Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP05/012631

International filing date: 01 July 2005 (01.07.2005)

Document type: Certified copy of priority document

Document details: Country/Office: JP
Number: 2004-197296
Filing date: 02 July 2004 (02.07.2004)

Date of receipt at the International Bureau: 22 July 2005 (22.07.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

PCT/JP2005/012631 PCT

日本国特許庁
JAPAN PATENT OFFICE

01.7.2005

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 Date of Application:	２００４年　７月　２日
出願番号 Application Number:	特願２００４－１９７２９６
パリ条約による外国への出願に用いる優先権の主張の基礎となる出願の国コードと出願番号 The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is	ＪＰ２００４－１９７２９６
出願人 Applicant(s):	日本電気株式会社 エヌイーシーコンピュータテクノ株式会社

２００５年　６月　９日

特許庁長官
Commissioner,
Japan Patent Office

小川洋

出証番号　出証特２００５－３０４９８７０

【書類名】　特許願
【整理番号】　33510057
【提出日】　平成16年 7月 2日
【あて先】　特許庁長官　殿
【国際特許分類】　G06F 13/00
G06F 15/00
【発明者】
【住所又は居所】　東京都港区芝五丁目７番１号　日本電気株式会社内
【氏名】　細見　岳生
【発明者】
【住所又は居所】　山梨県甲府市大津町１０８８－３　エヌイーシーコンピュータテクノ株式会社内
【氏名】　渡辺　佳晃
【特許出願人】
【識別番号】　000004237
【氏名又は名称】　日本電気株式会社
【特許出願人】
【識別番号】　000168285
【氏名又は名称】　エヌイーシーコンピュータテクノ株式会社
【代理人】
【識別番号】　100102864
【弁理士】
【氏名又は名称】　工藤　実
【手数料の表示】
【予納台帳番号】　053213
【納付金額】　16,000円
【提出物件の目録】
【物件名】　特許請求の範囲　1
【物件名】　明細書　1
【物件名】　図面　1
【物件名】　要約書　1
【包括委任状番号】　9715177
【包括委任状番号】　0016289

【書類名】特許請求の範囲
【請求項１】
ネットワークに接続された複数のプロセッサノードと、
前記ネットワークに接続された複数の入出力ノードと、
複数の入出力コントローラとを具備し、
前記複数のプロセッサノードの各々には、
複数のプロセッサと、
複数のデータを格納する主記憶部と、
前記複数のデータの各々に対してアクセス要求を受け付けることが可能なフリー状態情報が格納されたディレクトリと、
前記複数のプロセッサと前記主記憶部と前記ディレクトリとに接続されたメモリコントローラとが設けられ、
前記複数の入出力ノードの各々には、
ライトメッセージを発行する複数の入出力デバイスが設けられ、
前記複数の入出力ノードのうちの第１入出力ノードの前記複数の入出力デバイスによって１番目からM番目（Mは１以上の整数である）までのM個のデータに対するM個のライトメッセージが発行されたとき、前記複数の入出力コントローラのうちの第１入出力コントローラは、前記M個のそれぞれのデータに対するM個のライトトランザクションを開始し、前記M個のデータのうちのある第Ｉデータ（Ｉは、Ｉ＝１、２、…、Mを満たす整数の何れか）は、前記複数のプロセッサノードのうちの第１プロセッサノードをホームとするデータであり、第Ｉライトメッセージは前記第１プロセッサノードの前記主記憶部に格納された前記複数のデータのうちの第Ｉデータの値を第Ｉライトメッセージで指定される値に更新するための命令であり、前記第１入出力コントローラは、前記第Ｉライトトランザクションの処理として、第Ｉ書き込み要求メッセージを前記第１プロセッサノードに前記ネットワークを介して出力し、
前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ書き込み要求メッセージを受け取ったとき、前記第Ｉデータに対して、前記フリー状態情報に代えて、前記第Ｉデータに対するプロセッサや入出力デバイスからの読み出し要求や他の書き込み要求メッセージを受け付けることができないライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、前記第Ｉ書き込み要求メッセージに対して第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
前記第１入出力コントローラは、
前記第Ｉ書き込み許可メッセージを受け取ったときに第Ｉライトトランザクションの更新メッセージ発行処理を行い、
前記第１入出力コントローラは、前記更新メッセージ発行処理において、
第１から第Ｉ書き込み許可トランザクションまでのＩ個の書き込み許可メッセージを既に受け取っているか否かを検査し、
前記Ｉ個の書き込み許可メッセージをまだ受け取っていなければ、第Ｉ更新メッセージ発行処理を終了し、
前記Ｉ個の書き込み許可メッセージを既に受け取っているとき、
前記第Ｉライトメッセージで指定される値を含む第Ｉ更新メッセージを前記第１プロセッサノードに前記ネットワークを介して出力して第Ｉライトトランザクションを完了させ、（Ｉ＋１）がM以下であれば第（Ｉ＋１）ライトトランザクションの更新メッセージ発行処理を行い、（Ｉ＋１）がMより大きければ前記第Ｉ更新メッセージ発行処理を終了し、
前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ更新メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第１プロセッサノードの前記ディレクトリに格納すると共に、前記第１プロセッサノードの前記主記憶部に格納された前記第Ｉデータの値を前記第Ｉ更新メッセージで指定される値に更新する

マルチプロセッサシステム。
【請求項２】
請求項１に記載のマルチプロセッサシステムにおいて、
前記複数の入出力ノードには、それぞれ、前記複数の入出力コントローラが更に設けられ、
前記複数の入出力ノードの各々の入出力コントローラ（５０－ｊ）には、
前記複数の入出力ノードの各々の前記複数の入出力デバイスによって発行される前記Ｍ個のライトメッセージを調停するセレクタが更に設けられている
マルチプロセッサシステム。
【請求項３】
請求項１に記載のマルチプロセッサシステムにおいて、
前記ネットワークには、前記複数の入出力コントローラが更に接続され、
前記複数の入出力ノードの各々には、
前記複数の入出力ノードの各々の前記複数の入出力デバイスによって発行される前記Ｍ個のライトメッセージを調停するセレクタが更に設けられている
マルチプロセッサシステム。
【請求項４】
請求項１に記載のマルチプロセッサシステムにおいて、
前記複数のプロセッサノードには、それぞれ、前記複数の入出力コントローラが更に設けられ、
前記複数の入出力ノードの各々には、
前記複数の入出力ノードの各々の前記複数の入出力デバイスによって発行される前記Ｍ個のライトメッセージを調停するセレクタが更に設けられている
マルチプロセッサシステム。
【請求項５】
請求項１に記載のマルチプロセッサシステムにおいて、
前記複数のプロセッサノードと前記複数の入出力ノードとは、それぞれ複数のノードを構成し、
前記複数のノードの各々には、
前記複数のノードの各々の前記複数の入出力デバイスによって発行される前記Ｍ個のライトメッセージを調停するセレクタが更に設けられている
マルチプロセッサシステム。
【請求項６】
請求項１～５のいずれかに記載のマルチプロセッサシステムにおいて、
前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉデータに対して前記ライトロック状態情報が前記第１プロセッサノードの前記ディレクトリに格納されているときに、前記第Ｉデータに対する前記第Ｉ書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記第Ｉ書き込み要求メッセージに対して第Ｉ開放要求メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
前記第１入出力コントローラは、前記第Ｉ開放要求メッセージを受けて前記第Ｉ書き込み要求メッセージを前記第１プロセッサノードの前記メモリコントローラに前記ネットワークを介して出力すると共に、開放処理を行い、
前記第１入出力コントローラは、前記開放処理において、第Ｋライトトランザクション｛Ｋは、Ｋ＝Ｉ＋１、Ｉ＋２、…、Ｍを満たす整数であり、Ｉ＋１は、Ｉ＜（Ｉ＋１）＜Ｍを満たす整数であり、Ｉ＋２は、（Ｉ＋１）＜（Ｉ＋２）＜Ｍを満たす整数である｝の進捗を検査し、
未だ第Ｋ書き込み要求メッセージを発行していない場合、第Ｋ書き込み要求メッセージの発行を停止し、
既に第Ｋ書き込み要求メッセージを発行し第Ｋ書き込み許可メッセージを受け取っている場合、第Ｋ開放メッセージを前記第Ｋデータのホームである第２プロセッサノードに前

記ネットワークを介して出力し、
　既に第Ｋ書き込み要求メッセージを発行しまだ第Ｋ書きこみ許可メッセージを受け取っていない場合は、第Ｋ書き込み許可メッセージを受け取った時点で前記第Ｋ開放メッセージの発行を行い、
　前記第２プロセッサノードの前記メモリコントローラは、前記第Ｋ開放メッセージを受け取ったとき、前記第Ｋデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第２プロセッサノードの前記ディレクトリに格納する
　マルチプロセッサシステム。
【請求項７】
　請求項１～５のいずれかに記載のマルチプロセッサシステムにおいて、
　前記ディレクトリに格納される状態情報には、さらに割り込み可ライトロック状態情報とリクエストロック状態情報を含み、
　前記第１入出力コントローラが第Ｉ書き込み要求メッセージを発行する際に、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行し、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の強・弱書き込み要求メッセージ両方を受け付けることができないライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができない割り込み可ライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
　前記割り込み可ライトロック状態情報は、この状態にした弱書き込み要求メッセージを発行した入出力コントローラを特定する情報を含み、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報が前記ライトロック状態情報あるいはリクエストロック状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報が前記ライトロック状態に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報、前記ライトロック状態情報、あるいはリクエストロック状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに前記ネットワークを介して出力し、
　前記第１入出力コントローラは、前記第Ｉ不許可メッセージを受け取ったとき、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行し、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態

情報が前記割り込み可ライトロック状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラは、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記割り込み可ライトロック状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができないリクエストロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ再試行要求メッセージを前記ディレクトリに格納された入出力コントローラを特定する情報が指す第２入出力コントローラに宛てて前記ネットワークを介して出力し、

前記第Ｉ再試行要求メッセージを受け取った前記第２入出力コントローラは、再試行処理を行う

マルチプロセッサシステム。

【請求項８】

請求項７に記載のマルチプロセッサシステムにおいて、

前記第２入出力コントローラが行う再試行処理は、

第Ｉ開放メッセージを前記第１プロセッサノードの前記メモリコントローラに前記ネットワークを介して出力し、

前記第Ｉデータとアドレスを同じくするライトトランザクションで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ更新メッセージの発行を停止して、前記第Ｉ開放メッセージの発行後に、前記第１プロセッサノードに宛てて書き込み要求メッセージを発行し、

前記第Ｉ開放メッセージを受けた前記第１プロセッサノードの前記メモリコントローラは、前記リクエストロック状態情報に代えてライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに宛てて前記ネットワークを介して出力する

マルチプロセッサシステム。

【請求項９】

複数のプロセッサノードと、複数の入出力ノードと、複数の入出力コントローラとを具備し、

前記複数のプロセッサノードの各々には、複数のプロセッサと、複数のデータを格納する主記憶部と、前記複数のデータの各々に対してアクセス要求を受け付けることが可能なフリー状態情報が格納されたディレクトリと、前記複数のプロセッサと前記主記憶部と前記ディレクトリとに接続されたメモリコントローラとが設けられ、

前記複数の入出力ノードの各々には、ライトメッセージを発行する複数の入出力デバイスが設けられた、

マルチプロセッサシステムに適用されるメモリアクセス処理方法であって、

前記メモリアクセス処理方法は、

前記複数の入出力ノードのうちの第１入出力ノードの前記複数の入出力デバイスによって１番目からＭ番目（Ｍは１以上の整数である）までのＭ個のデータに対するＭ個のライトメッセージが発行されたとき、前記複数の入出力コントローラのうちの第１入出力コントローラが、前記Ｍ個のそれぞれのデータに対するＭ個のライトトランザクションを開始し、前記Ｍ個のデータのうちのある第Ｉデータ（Ｉは、Ｉ＝１、２、…、Ｍを満たす整数の何れか）は、前記複数のプロセッサノードのうちの第１プロセッサノードをホームとするデータであり、第Ｉライトメッセージは前記第１プロセッサノードの前記主記憶部に格納された前記複数のデータのうちの第Ｉデータの値を第Ｉライトメッセージで指定される値に更新するための命令であり、前記第１入出力コントローラが、前記第Ｉライトトランザクションの処理として、第Ｉ書き込み要求メッセージを前記第１プロセッサノードに出力するステップと、

前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ書き込み要求メッセージを受け取ったとき、前記第Ｉデータに対して、前記フリー状態情報に代えて、前記第Ｉデータに対するプロセッサや入出力デバイスからの読み出し要求や他の書き込み要求メ

ッセージを受け付けることができないライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、前記第Ｉ書き込み要求メッセージに対して第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、
前記第１入出力コントローラが、前記第Ｉ書き込み許可メッセージを受け取ったとき、第Ｉライトトランザクションの処理として更新メッセージ発行処理を行うステップとを含み、
前記第１入出力コントローラが行う前記更新メッセージ発行処理は、
第１から第Ｉ書き込み許可メッセージまでのＩ個の書き込み許可メッセージを既に受け取っているか否かを検査するステップと、
前記Ｉ個の書き込み許可メッセージをまだ受け取っていなければ、前記第Ｉ更新メッセージ発行処理を終了するステップと、
前記Ｉ個の書き込み許可メッセージを既に受け取っていれば、前記第Ｉライトメッセージで指定される値を含む第Ｉ更新メッセージを前記第１プロセッサノードに出力して第Ｉライトトランザクションを完了させ、（Ｉ＋１）がM以下であれば第（Ｉ＋１）ライトトランザクションの更新メッセージ発行処理を行い、（Ｉ＋１）がMより大きければ前記第Ｉ更新メッセージ発行処理を終了するステップとを含み、
前記メモリアクセス処理方法は、更に、
前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ更新メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第１プロセッサノードの前記ディレクトリに格納すると共に、前記第１プロセッサノードの前記主記憶部に格納された前記第Ｉデータの値を前記第Ｉ更新メッセージで指定される値に更新するステップと
を含むメモリアクセス処理方法。
【請求項１０】
請求項９に記載のメモリアクセス処理方法において、
更に、
前記複数の入出力ノードの各々の前記複数の入出力デバイスによって発行される前記M個のライトメッセージを調停するステップ
を含むメモリアクセス処理方法。
【請求項１１】
請求項９又は１０に記載のメモリアクセス処理方法において、
更に、
前記第Ｉデータに対して前記ライトロック状態情報が前記第１プロセッサノードの前記ディレクトリに格納されているときに、前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉデータに対する前記第Ｉ書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記第Ｉ書き込み要求メッセージに対して第Ｉ開放要求メッセージを前記第１入出力コントローラに出力するステップと、
前記第１入出力コントローラが、前記第Ｉ開放要求メッセージを受けて前記第Ｉ書き込み要求メッセージを前記第１プロセッサノードの前記メモリコントローラに出力すると共に、開放処理を行なうステップとを含み、
前記開放処理を行なうステップは、
前記第Ｉデータに後続する第Ｋライトトランザクション｛Kは、Ｋ＝Ｉ＋１、Ｉ＋２、…、Mを満たす整数であり、Ｉ＋１は、Ｉ＜（Ｉ＋１）＜Mを満たす整数であり、Ｉ＋２は、（Ｉ＋１）＜（Ｉ＋２）＜Mを満たす整数である｝の進捗を検査し、未だ第Ｋ書き込み要求メッセージを発行していない場合、第Ｋ書き込み要求メッセージの発行を停止するステップと、
既に第Ｋ書き込み要求メッセージを発行し第Ｋ書き込み許可メッセージを受け取っている場合、第Ｋ開放メッセージを前記第Ｋデータのホームである第２プロセッサノードの前記メモリコントローラに出力するステップと、
既に第Ｋ書き込み要求メッセージを発行しまだ第Ｋ書き込み許可メッセージを受け取っ

ていない場合は、第Ｋ書き込み許可メッセージを受け取った時点で前記第Ｋ開放メッセージの発行を行うステップとを含み、
　前記メモリアクセス処理方法は、更に、
　前記第２プロセッサノードの前記メモリコントローラが、前記第Ｋ開放メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第２プロセッサノードの前記ディレクトリに格納するステップを
含むメモリアクセス処理方法。
【請求項１２】
　請求項９又は１０に記載のメモリアクセス処理方法において、
　前記ディレクトリに格納される状態情報には、さらに割り込み可ライトロック状態情報とリクエストロック状態情報を含み、
　前記メモリアクセス処理方法は、更に、
　前記第１入出力コントローラが第Ｉ書き込み要求メッセージを発行する際に、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行するステップと、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の強・弱書き込み要求メッセージ両方を受け付けることができないライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができない割り込み可ライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、
　前記割り込み可ライトロック状態情報は、この状態にした弱書き込み要求メッセージを発行した入出力コントローラを特定する情報を含み、
　前記メモリアクセス処理方法は、更に、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報が前記ライトロック状態情報あるいはリクエストロック状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに出力するステップと、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報が前記ライトロック状態に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報、前記ライトロック状態情報、あるいはリクエストロック状態情報であるときに、前記第１プロセッサノードの前記メモリコントローラが、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに出力するステップと、
　前記第１入出力コントローラが、前記第Ｉ不許可メッセージを受け取ったとき、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行するステップと、
　前記第１プロセッサノードの前記ディレクトリに格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報であるときに、前記第１プロセッサノードの

前記メモリコントローラが、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記割り込み可ライトロック状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができないリクエストロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ再試行要求メッセージを前記ディレクトリに格納された入出力コントローラを特定する情報が指す第２入出力コントローラに宛てて出力するステップと、

前記第Ｉ再試行要求メッセージを受け取った前記第２入出力コントローラが、再試行処理を行うステップと

を含むメモリアクセス処理方法。

【請求項１３】

請求項１２に記載のメモリアクセス処理方法において、

前記第２入出力コントローラが再試行処理を行うステップは、

第Ｉ開放メッセージを前記第１プロセッサノードの前記メモリコントローラに出力するステップと、

前記第Ｉデータとアドレスを同じくするライトトランザクションで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ更新メッセージの発行を停止して、前記第Ｉ開放メッセージの発行後に、前記第１プロセッサノードに宛てて書き込み要求メッセージを発行するステップと、

前記第Ｉ開放メッセージを受けた前記第１プロセッサノードの前記メモリコントローラは、前記リクエストロック状態情報に代えてライトロック状態情報を前記第１プロセッサノードの前記ディレクトリに格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに宛てて出力するステップと

を含むメモリアクセス処理方法。

【書類名】明細書
【発明の名称】マルチプロセッサシステムおよびメモリアクセス処理方法
【技術分野】
【０００１】
本発明は、メモリを共有する密結合型のマルチプロセッサシステムに関し、入出力デバイスからのメモリアクセスを処理するマルチプロセッサシステムおよびメモリアクセス処理方法に関する。
【背景技術】
【０００２】
ＰＣＩバス仕様リビジョン２．１等の規定では、ＰＣＩバス上の入出力デバイスを要求元にしたライトメッセージは順序を保障しなければならないという制約がある。即ち、先行のメッセージが完了してから後続のメッセージが完了することを保障しなければならない。
【０００３】
図１はディレクトリ方式によってデータの一貫性を維持するマルチプロセッサシステムの構成を示す図である。マルチプロセッサシステムは、複数のプロセッサノード１０１－１～１０１－ｍ（ｍは１以上の整数）と、複数の入出力ノード１０３－１～１０３－ｎ（ｎは１以上の整数）とを具備している。複数のプロセッサノード１０１－１～１０１－ｍと複数の入出力ノード１０３－１～１０３－ｎとは、ネットワーク１０２に接続され、外部からのクロックに応じて動作する。プロセッサノード１０１－ｉ（ｉ＝１、２、…、ｍ）には、プロセッサ１１０－ｉ－１、１１０－ｉ－２と、ディレクトリ１２０－ｉと、主記憶部（メモリ）１３０－ｉと、メモリコントローラ１４０－ｉとが設けられている。入出力ノード１０３－ｊ（ｊ＝１、２、…、ｎ）は、入出力コントローラ１５０－ｊと、外部からの命令によりメッセージを発行する複数の入出力デバイス１６０－ｊ－１、１６０－ｊ－２とが設けられている。
主記憶部１３０－ｉには、複数のデータが格納されている。複数のデータの各々は、その内容を表す値を含んでいる。
【０００４】
図１に示すようなマルチプロセッサシステムにおいて前記ライトメッセージの順序制約を満たす従来技術を紹介する。特許文献１の［発明が解決しようとする課題］に記載された技術(以降、従来技術１)は、先行するライトメッセージの完了が保障されるまで後続するライトメッセージの発行を留めることでこの制約を満たすというものである。
【０００５】
図２を参照しながら、順序制約のあるライトメッセージが連続して発行された場合の従来技術１の動作を説明する。図２は、入出力デバイス１６０－１－１が、データＡ、Ｂ、Ｃに対するライトメッセージをそれぞれステップ１、２、３で発行した場合の動作を示している。ここで、データＡおよびデータＢはプロセッサノード１－１をホームとし、データＣはプロセッサノード１－２をホームとするデータであるとする。また、１ステップは１クロックに対応する。
【０００６】
入出力コントローラ１５０－１は、ステップ２にて、入出力デバイス１６０－１－１からのライトメッセージとしてライトＡメッセージを受け取る。このとき、入出力コントローラ１５０－１は、ステップ３にて、ライトＡメッセージで指定される値を含む更新Ａメッセージをホームのメモリコントローラ１４０－１に宛ててネットワーク１０２に出力する。
メモリコントローラ１４０－１は、ステップ４にて、入出力コントローラ１５０－１からの更新Ａメッセージを受けて、主記憶部１３０－１に格納されたデータの値を更新Ａメッセージで指定される値に更新する。このとき、メモリコントローラ１４０－１は、ステップ５にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ａメッセージを出力する。

入出力コントローラ１５０－１は、ステップ６にて、メモリコントローラ１４０－１からの完了Ａメッセージを受け取り、先行するライトＡが完了したことを認識する。

【０００７】

入出力コントローラ１５０－１は、ステップ３にて、第２ライトメッセージとしてライトＢメッセージを受け取る。このとき、入出力コントローラ１５０－１は、先行するライトＡメッセージが完了するステップ６までライトＢメッセージを留め置き、ステップ７にて、ライトＢメッセージで指定される値を含む更新Ｂメッセージをメモリコントローラ１４０－１に宛ててネットワーク１０２に出力する。

メモリコントローラ１４０－１は、ステップ８にて、入出力コントローラ１５０－１からの更新Ｂメッセージを受けて、主記憶部１３０－１に格納されたデータの値を更新Ｂメッセージで指定される値に更新する。このとき、メモリコントローラ１４０－１は、ステップ９にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ｂメッセージを出力する。

入出力コントローラ１５０－１は、ステップ１０にて、メモリコントローラ１４０－１からの完了Ｂメッセージを受け取り、先行するライトＢが完了したことを認識する。

【０００８】

入出力コントローラ１５０－１は、ステップ４にて、第３ライトメッセージとしてライトＣメッセージを受け取る。このとき、入出力コントローラ１５０－１は、先行するライトＡメッセージおよびライトＢメッセージが両方とも完了するステップ１０までライトＣメッセージを留め置き、ステップ１１にて、ライトＣメッセージで指定される値を含む更新Ｃメッセージをメモリコントロー１４０－２に宛ててネットワーク１０２に出力する。

メモリコントローラ１４０－２は、ステップ１２にて、入出力コントローラ１５０－１からの更新Ｃメッセージを受けて、主記憶部１３０－２に格納されたデータの値を更新Ｃメッセージで指定される値に更新する。このとき、メモリコントローラ１４０－２は、ステップ１３にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ｃメッセージを出力する。

入出力コントローラ１５０－１は、ステップ１４にて、メモリコントローラ１４０－２からの完了Ｃメッセージを受け取り、ライトＣが完了したことを認識する。

【０００９】

このように、入出力コントローラ１５０－ｊは、先行するライトメッセージが完了してから次のライトメッセージを発行することで、入出力デバイス１６０－ｊ－１、６０－ｊ－２が発行した複数のライトメッセージの順序を保障することができる。しかし、３つのライトメッセージを処理するのに１４ステップを要する。

【００１０】

このライトメッセージを処理するのに要する時間が長く、性能が劣化する問題を解決する従来技術（従来技術２）が特許文献１に記載されている。この技術は、同一プロセッサノードを宛先とするライトメッセージの連続発行、即ち先行するメッセージの完了が保障される前に後続するメッセージを発行することを可能とするものである。

【００１１】

図３を参照しながら、従来技術２の動作を説明する。

【００１２】

入出力コントローラ１５０－１は、ステップ２にて、第１ライトメッセージとしてライトＡメッセージを受け取り、ライトＡメッセージがメモリコントローラ１４０－１をホームとするライトであることを認識する。このとき、入出力コントローラ１５０－１は、ステップ３にて、ライトＡメッセージで指定される値を含む更新Ａメッセージをホームのメモリコントロー１４０－１に宛ててネットワーク１０２に出力する。

メモリコントローラ１４０－１は、ステップ４にて、入出力コントローラ１５０－１からの更新Ａメッセージを受けて主記憶部１３０－１のデータを更新する。このとき、メモリコントローラ１４０－１は、ステップ５にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ａメッセージを出力する。

入出力コントローラ１５０－１は、ステップ6にて、メモリコントローラ１４０－１からの完了Ａメッセージを受け取り、先行するライトＡメッセージが完了したことを認識する。

【００１３】

入出力コントローラ１５０－１は、ステップ3にて、第2ライトメッセージとしてライトＢメッセージを受け取り、ライトＢメッセージがメモリコントローラ１４０－１をホームとするライトであることを認識し、先行するライトＡメッセージと同じホーム（メモリコントローラ１４０－１）であることを認識する。このとき、入出力コントローラ１５０－１は、ライトＡメッセージの完了を待つことなく、ステップ4にて、ライトＢメッセージで指定される値を含む更新Ｂメッセージをホームのメモリコントロー１４０－１に宛ててネットワーク１０２に出力する。

メモリコントローラ１４０－１は、ステップ5にて、入出力コントローラ１５０－１からの更新Ｂメッセージを受けて主記憶部１３０－１のデータを更新する。このとき、メモリコントローラ１４０－１は、ステップ6にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ｂメッセージを出力する。

入出力コントローラ１５０－１は、ステップ7にて、メモリコントローラ１４０－１からのステップ7で完了Ｂメッセージを受け取る。

【００１４】

入出力コントローラ１５０－１は、ステップ4にて、第3ライトメッセージとしてライトＣメッセージを受け取り、ライトＣメッセージがメモリコントローラ１４０－２をホームとするライトであることを認識し、先行するライトＡメッセージ、ライトＢメッセージとはホーム（メモリコントローラ１４０－１）が異なることを認識する。このとき、入出力コントローラ１５０－１は、両ライトメッセージ（ライトＡメッセージ、ライトＢメッセージ）とも完了するステップ7までライトＣメッセージを留め置き、ステップ8にて、ライトＣメッセージで指定される値を含む更新Ｃメッセージをメモリコントローラ１４０－２に宛ててネットワーク１０２に出力する。

メモリコントローラ１４０－２は、ステップ9にて、入出力コントローラ１５０－１からの更新Ｃメッセージを受けて主記憶部１３０－２のデータを更新する。このとき、メモリコントローラ１４０－２は、ステップ１０にて、入出力コントローラ１５０－１に宛ててネットワーク１０２に完了Ｃメッセージを出力する。

入出力コントローラ１５０－１は、ステップ１１にて、メモリコントローラ１４０－１からの完了Ｃメッセージを受け取り、ライトＣが完了したことを認識する。

【００１５】

このように、入出力コントローラ１５０－ｊは、ホームを同じにするライトメッセージについては連続して発行し、ホームが異なるライトメッセージについては先行するライトメッセージが完了してから発行する。ネットワーク１０２は2点間のメッセージの順序を保障するので、上記例ではライトＡメッセージとライトＢメッセージはその順でメモリコントローラ１４０－１に到着することが保障される。そのため、ライトＡメッセージを追い越してライトＢメッセージが先にメモリコントローラ１４０－１で処理されることはなく順序を保障することができる。

【００１６】

しかし、この従来技術2でも3つのライトメッセージを処理するのに１１ステップを要する。

【００１７】

【特許文献１】特開２００１－２１６２５９号公報

【発明の開示】

【発明が解決しようとする課題】

【００１８】

上記の従来技術１、従来技術２では、入出力コントローラが、異なるプロセッサノードを宛先とする入出力デバイスからの複数のライトメッセージを連続して処理することがで

きない。このため、入出力コントローラがライトメッセージの処理を行う場合に要する時間は、長くなってしまう。

【００１９】

本発明の課題は、入出力コントローラが、異なるプロセッサノードを宛先とする入出力デバイスからの複数のライトメッセージを連続して処理することができるマルチプロセッサシステムを提供することにある。

本発明の他の課題は、入出力コントローラがライトメッセージの処理を行う場合に要する時間を短くすることができるマルチプロセッサシステムを提供することにある。

【課題を解決するための手段】

【００２０】

以下に、［発明を実施するための最良の形態］で使用する番号・符号を用いて、課題を解決するための手段を説明する。これらの番号・符号は、［特許請求の範囲］の記載と［発明を実施するための最良の形態］の記載との対応関係を明らかにするために付加されたものであるが、［特許請求の範囲］に記載されている発明の技術的範囲の解釈に用いてはならない。

【００２１】

本発明のマルチプロセッサシステムは、ネットワーク（２）に接続された複数のプロセッサノード（１－１～１－ｍ）（ｍは１以上の整数である）と、前記ネットワーク（２）に接続された複数の入出力ノード（３－１～３－ｎ）（ｎは１以上の整数である）と、複数の入出力コントローラとを具備している。

前記複数のプロセッサノードの各々（１－ｉ）（ｉ＝１、２、…、ｍ）には、複数のプロセッサ（１０－ｉ－１、１０－ｉ－２）と、複数のデータを格納する主記憶部（３０－ｉ）と、前記複数のデータの各々に対してアクセス要求を受け付けることが可能なフリー状態情報が格納されたディレクトリ（２０－ｉ）と、前記複数のプロセッサ（１０－ｉ－１、１０－ｉ－２）と前記主記憶部（３０－ｉ）と前記ディレクトリ（２０－ｉ）とに接続されたメモリコントローラ（４０－ｉ）とが設けられている。

前記複数の入出力ノードの各々（３－ｊ）（ｊ＝１、２、…、ｎ）には、ライトメッセージを発行する複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）が設けられている。

前記複数の入出力ノード（３－１～３－ｎ）のうちの第１入出力ノード（３－１）の前記複数の入出力デバイス（６０－１－１、６０－１－２）によって１番目からＭ番目（Ｍは１以上の整数である）までのＭ個のデータに対するＭ個のライトメッセージが発行されたとき、前記複数の入出力コントローラのうちの第１入出力コントローラは、前記Ｍ個のそれぞれのデータに対するＭ個のライトトランザクションを開始する。前記Ｍ個のデータのうちのある第Ｉデータ（Ｉは、Ｉ＝１、２、…、Ｍを満たす整数の何れか）は、前記複数のプロセッサノード（１－１～１－ｍ）のうちの第１プロセッサノード（１－ｉ）をホームとするデータである。第Ｉライトトランザクションは前記第１プロセッサノード（１－ｉ）の前記主記憶部（３０－ｉ）に格納された前記複数のデータのうちの第Ｉデータの値を第Ｉライトトランザクションで指定される値に更新するための命令である。前記第１入出力コントローラは、前記第Ｉライトトランザクションの処理として、第Ｉ書き込み要求メッセージを前記第１プロセッサノード（１－ｉ）に前記ネットワーク（２）を介して出力する。

前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ書き込み要求メッセージを受け取ったとき、前記第Ｉデータに対して、前記フリー状態情報に代えて、前記第Ｉデータに対するプロセッサや入出力デバイスからの読み出し要求や他の書き込み要求メッセージを受け付けることができないライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、前記第Ｉ書き込み要求メッセージに対して第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記第Ｉ書き込み許可メッセージを受け取ったときに第Ｉライトトランザクションの更

新メッセージ発行処理を行う。前記第１入出力コントローラは、前記更新メッセージ発行処理において、第１から第Ｉ書き込み許可トランザクションまでのＩ個の書き込み許可メッセージを既に受け取っているか否かを検査し、前記Ｉ個の書き込み許可メッセージをまだ受け取っていなければ、第Ｉ更新メッセージ発行処理を終了し、前記Ｉ個の書き込み許可メッセージを既に受け取っているとき、前記第Ｉライトメッセージで指定される値を含む第Ｉ更新トランザクションを前記第１プロセッサノードに前記ネットワーク（２）を介して出力して第Ｉライトトランザクションを完了させ、（Ｉ＋１）がM以下であれば第（Ｉ＋１）ライトトランザクションの更新メッセージ発行処理を行い、（Ｉ＋１）がMより大きければ前記第Ｉ更新メッセージ発行処理を終了する。

前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ更新メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納すると共に、前記第１プロセッサノード（１－ｉ）の前記主記憶部（３０－ｉ）に格納された前記第Ｉデータの値を前記第Ｉ更新メッセージで指定される値に更新する。

【００２２】

本発明のマルチプロセッサシステムにおいて、前記複数の入出力ノード（３－１～３－ｎ）には、それぞれ、前記複数の入出力コントローラ（５０－１～５０－ｎ）が更に設けられている。

前記複数の入出力ノードの各々（３－ｊ）（ｊ＝１、２、…、ｎ）の入出力コントローラ（５０－ｊ）には、前記複数の入出力ノードの各々（３－ｊ）の前記複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）によって発行される前記M個のライトメッセージを調停するセレクタ（７１－ｊ）が更に設けられている。

【００２３】

本発明のマルチプロセッサシステムにおいて、前記ネットワーク（２）には、前記複数の入出力コントローラ（５２－１～５２－ｎ）が更に接続されている。

前記複数の入出力ノードの各々（３－ｊ）（ｊ＝１、２、…、ｎ）には、前記複数の入出力ノードの各々（３－ｊ）の前記複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）によって発行される前記M個のライトメッセージを調停するセレクタ（５１－ｊ）が更に設けられている。

【００２４】

本発明のマルチプロセッサシステムにおいて、前記複数のプロセッサノード（１－１～１－ｍ）には、それぞれ、前記複数の入出力コントローラ（５２－１～５２－ｍ）（図示しない）が更に設けられている。

前記複数の入出力ノードの各々（３－ｊ）（ｊ＝１、２、…、ｎ）には、前記複数の入出力ノードの各々（３－ｊ）の前記複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）によって発行される前記M個のライトメッセージを調停するセレクタ（５１－ｊ）が更に設けられている。

【００２５】

本発明のマルチプロセッサシステムにおいて、前記複数のプロセッサノード（１－１～１－ｍ）と前記複数の入出力ノード（３－１～３－ｎ）とは、それぞれ複数のノード（図示しない）（ｍ＝ｎ）を構成する。

前記複数のノード（図示しない）の各々には、前記複数のノード（図示しない）の各々の前記複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）によって発行される前記M個のライトメッセージを調停するセレクタ（５１－ｊ）が更に設けられている。

【００２６】

本発明のマルチプロセッサシステムにおいて、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉデータに対して前記ライトロック状態情報が前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されているときに、前記第Ｉデータに対する前記第Ｉ書き込み要求メッセージを前記第１入出

力コントローラから受け取った場合、前記第Ｉ書き込み要求メッセージに対して第Ｉ開放要求メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記第１入出力コントローラは、前記第Ｉ開放要求メッセージを受けて前記第Ｉ書き込み要求メッセージを前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）に前記ネットワーク（２）を介して出力すると共に、開放処理を行なう。

前記第１入出力コントローラは、前記開放処理において、第Ｋライトトランザクション｛Ｋは、Ｋ＝Ｉ＋１、Ｉ＋２、…、Ｍを満たす整数であり、Ｉ＋１は、Ｉ＜（Ｉ＋１）＜Ｍを満たす整数であり、Ｉ＋２は、（Ｉ＋１）＜（Ｉ＋２）＜Ｍを満たす整数である｝の進捗を検査し、未だ第Ｋ書き込み要求メッセージを発行していない場合、第Ｋ書き込み要求メッセージの発行を停止する。既に第Ｋ書き込み要求トランザクションを発行し第Ｋ書き込み許可トランザクションを受け取っている場合、前記第１入出力コントローラは、第Ｋ開放トランザクションを前記第Ｋデータのホームである第２プロセッサノード（１－ｋ）（ｋ＝１、２、…、ｍ）に前記ネットワーク（２）を介して出力する。既に第Ｋ書き込み要求メッセージを発行しまだ第Ｋ書きこみ許可メッセージを受け取っていない場合は、前記第１入出力コントローラは、第Ｋ書き込み許可メッセージを受け取った時点で前記第Ｋ開放メッセージの発行を行う。

前記第２プロセッサノード（１－ｋ）の前記メモリコントローラ（４０－ｋ）は、前記第Ｋ開放メッセージを受け取ったとき、前記第Ｋデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第２プロセッサノード（１－ｋ）の前記ディレクトリ（２０－ｋ）に格納する。

【００２７】

前記ディレクトリ（２０－ｉ）に格納される状態情報には、さらに割り込み可ライトロック状態情報とリクエストロック状態情報を含んでいる。

前記第１入出力コントローラが第Ｉ書き込み要求メッセージを発行する際に、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行する。

前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の強・弱書き込み要求メッセージ両方を受け付けることができないライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができない割り込み可ライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記割り込み可ライトロック状態情報は、この状態にした弱書き込み要求メッセージを発行した入出力コントローラを特定する情報を含んでいる。

前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記ライトロック状態情報あるいはリクエストロック状態

情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記ライトロック状態に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報、前記ライトロック状態情報、あるいはリクエストロック状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに前記ネットワーク（２）を介して出力する。

前記第１入出力コントローラは、前記第Ｉ不許可メッセージを受け取ったとき、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行する。

前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記割り込み可ライトロック状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができないリクエストロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ再試行要求メッセージを前記ディレクトリ（２０－ｉ）に格納された入出力コントローラを特定する情報が指す第２入出力コントローラに宛てて前記ネットワーク（２）を介して出力する。

前記第Ｉ再試行要求メッセージを受け取った前記第２入出力コントローラは、再試行処理を行う。

【００２８】

本発明のマルチプロセッサシステムにおいて、前記第２入出力コントローラは以下のように再試行処理を行う。

第Ｉ開放メッセージを前記第１プロセッサノード（２０－ｉ）の前記メモリコントローラ（４０－ｉ）に前記ネットワーク（２）を介して出力する。

前記第Ｉデータとアドレスを同じくするライトメッセージで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ更新メッセージの発行を停止して、前記第Ｉ開放メッセージの発行後に、前記第１プロセッサノード（２０－ｉ）に宛てて書き込み要求メッセージを発行する。

前記第Ｉ開放メッセージを受けた前記第１プロセッサノード（２０－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記リクエストロック状態情報に代えてライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに宛てて前記ネットワーク（２）を介して出力する。

【００２９】

本発明のメモリアクセス処理方法は、マルチプロセッサシステムに適用される。マルチプロセッサシステムは、複数のプロセッサノード（１－１～１－ｍ）（ｍは１以上の整数である）と、複数の入出力ノード（３－１～３－ｎ）（ｎは１以上の整数である）と、複数の入出力コントローラとを具備している。

前記複数のプロセッサノードの各々（１－ｉ）（ｉ＝１、２、…、ｍ）には、複数のプロセッサ（１０－ｉ－１、１０－ｉ－２）と、複数のデータを格納する主記憶部（３０－ｉ）と、前記複数のデータの各々に対してアクセス要求を受け付けることが可能なフリー状態情報が格納されたディレクトリ（２０－ｉ）と、前記複数のプロセッサ（１０－ｉ－

１、１０－ｉ－2）と前記主記憶部（３０－ｉ）と前記ディレクトリ（２０－ｉ）とに接続されたメモリコントローラ（４０－ｉ）とが設けられている。前記複数の入出力ノードの各々（３－ｊ）（ｊ＝１、２、…、ｎ）には、ライトメッセージを発行する複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）が設けられている。

本発明のメモリアクセス処理方法は、前記複数の入出力ノード（３－１～３－ｎ）のうちの第１入出力ノード（３－１）の前記複数の入出力デバイス（６０－１－１、６０－１－２）によって１番目からＭ番目（Ｍは１以上の整数である）までのＭ個のデータに対するＭ個のライトメッセージが発行されたとき、前記複数の入出力コントローラのうちの第１入出力コントローラが、前記Ｍ個のそれぞれのデータに対するＭ個のライトトランザクションを開始し、前記Ｍ個のデータのうちのある第Ｉデータ（Ｉは、Ｉ＝１、２、…、Ｍを満たす整数の何れか）は、前記複数のプロセッサノード（１－１～１－ｍ）のうちの第１プロセッサノード（１－ｉ）をホームとするデータであり、第Ｉライトメッセージは前記第１プロセッサノード（１－ｉ）の前記主記憶部（３０－ｉ）に格納された前記複数のデータのうちの第Ｉデータの値を第Ｉライトメッセージで指定される値に更新するための命令であり、前記第１入出力コントローラが、前記第Ｉライトトランザクションの処理として、第Ｉ書き込み要求メッセージを前記第１プロセッサノード（１－ｉ）に出力するステップと、

前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ書き込み要求メッセージを受け取ったとき、前記第Ｉデータに対して、前記フリー状態情報に代えて、前記第Ｉデータに対するプロセッサや入出力デバイスからの読み出し要求や他の書き込み要求メッセージを受け付けることができないライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、前記第Ｉ書き込み要求メッセージに対して第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、前記第１入出力コントローラが、前記第Ｉ書き込み許可メッセージを受け取ったとき、第Ｉライトトランザクションの処理として更新メッセージ発行処理を行うステップとを含む。

前記第１入出力コントローラが行う前記更新メッセージ発行処理は、第１から第Ｉ書き込み許可メッセージまでのＩ個の書き込み許可メッセージを既に受け取っているか否かを検査するステップと、前記Ｉ個の書き込み許可メッセージをまだ受け取っていなければ、前記第Ｉ更新メッセージ発行処理を終了するステップと、前記Ｉ個の書き込み許可メッセージを既に受け取っていれば、前記第Ｉライトメッセージで指定される値を含む第Ｉ更新メッセージを前記第１プロセッサノードに出力して第Ｉライトトランザクションを完了させ、（Ｉ＋１）がＭ以下であれば第（Ｉ＋１）ライトトランザクションの更新メッセージ発行処理を行い、（Ｉ＋１）がＭより大きければ前記第Ｉ更新メッセージ発行処理を終了するステップとを含む。

前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ更新メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納すると共に、前記第１プロセッサノード（１－ｉ）の前記主記憶部（３０－ｉ）に格納された前記第Ｉデータの値を前記第Ｉ更新メッセージで指定される値に更新するステップとを含んでいる。

【００３０】

本発明のメモリアクセス処理方法は、更に、前記複数の入出力ノードの各々（３－ｊ）の前記複数の入出力デバイス（６０－ｊ－１、６０－ｊ－２）によって発行される前記Ｍ個のライトメッセージを調停するステップを含んでいる。

【００３１】

本発明のメモリアクセス処理方法は、更に、前記第Ｉデータに対して前記ライトロック状態情報が前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されているときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉデータに対する前記第Ｉ書き込み要求メッセージを前記第１入出力

コントローラから受け取った場合、前記第Ｉ書き込み要求メッセージに対して第Ｉ開放要求メッセージを前記第１入出力コントローラに出力するステップと、前記第１入出力コントローラが、前記第Ｉ開放要求メッセージを受けて前記第Ｉ書き込み要求メッセージを前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）に出力すると共に、開放処理を行なうステップとを含んでいる。

前記開放処理を行なうステップは、前記第Ｉデータに後続する第Ｋライトトランザクション｛Ｋは、Ｋ＝Ｉ＋１、Ｉ＋２、…、Ｍを満たす整数であり、Ｉ＋１は、Ｉ＜（Ｉ＋１）＜Ｍを満たす整数であり、Ｉ＋２は、（Ｉ＋１）＜（Ｉ＋２）＜Ｍを満たす整数である｝の進捗を検査し、未だ第Ｋ書き込み要求メッセージを発行していない場合、第Ｋ書き込み要求メッセージの発行を停止するステップと、既に第Ｋ書き込み要求メッセージを発行し第Ｋ書き込み許可メッセージを受け取っている場合、第Ｋ開放メッセージを前記第Ｋデータのホームである第２プロセッサノード（１－ｋ）の前記メモリコントローラ（４０－ｋ）に出力するステップと、既に第Ｋ書き込み要求メッセージを発行しまだ第Ｋ書き込み許可メッセージを受け取っていない場合は、第Ｋ書き込み許可メッセージを受け取った時点で前記第Ｋ開放メッセージの発行を行うステップとを含む。

前記メモリアクセス処理方法は、更に、前記第２プロセッサノード（１－ｋ）の前記メモリコントローラ（４０－ｋ）が、前記第Ｋ開放メッセージを受け取ったとき、前記第Ｉデータに対して前記ライトロック状態情報に代えて前記フリー状態情報を前記第２プロセッサノード（１－ｋ）の前記ディレクトリ（２０－ｋ）に格納するステップを含んでいる。

【００３２】

本発明のメモリアクセス処理方法において、前記ディレクトリ（２０－ｉ）に格納される状態情報には、さらに割り込み可ライトロック状態情報とリクエストロック状態情報を含んでいる。

本発明のメモリアクセス処理方法は、更に、前記第１入出力コントローラが第Ｉ書き込み要求メッセージを発行する際に、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行するステップと、前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の強・弱書き込み要求メッセージ両方を受け付けることができないライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報がフリー状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができない割り込み可ライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに出力するステップと、前記割り込み可ライトロック状態情報は、この状態にした弱書き込み要求メッセージを発行した入出力コントローラを特定する情報を含んでいる。

本発明のメモリアクセス処理方法は、更に、前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記ライトロック状態情報あるいはリクエストロック状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ強書き込み要求メッセ

ージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに出力するステップと、前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記ライトロック状態に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報、前記ライトロック状態情報、あるいはリクエストロック状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ弱書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、第Ｉ不許可メッセージを前記第１入出力コントローラに出力するステップと、前記第１入出力コントローラが、前記第Ｉ不許可メッセージを受け取ったとき、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行するステップと、前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納されている前記第Ｉデータの状態情報が前記割り込み可ライトロック状態情報であるときに、前記第１プロセッサノード（１－ｉ）の前記メモリコントローラ（４０－ｉ）が、前記第Ｉ強書き込み要求メッセージを前記第１入出力コントローラから受け取った場合、前記割り込み可ライトロック状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができないリクエストロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ再試行要求メッセージを前記ディレクトリ（２０－ｉ）に格納された入出力コントローラを特定する情報が指す第２入出力コントローラに宛てて出力するステップと、前記第Ｉ再試行要求メッセージを受け取った前記第２入出力コントローラが、再試行処理を行うステップとを含んでいる。

【００３３】

本発明のメモリアクセス処理方法において、前記第２入出力コントローラが再試行処理を行うステップは、第Ｉ開放メッセージを前記第１プロセッサノード（２０－ｉ）の前記メモリコントローラ（４０－ｉ）に出力するステップと、前記第Ｉデータとアドレスを同じくするライトメッセージで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ更新メッセージの発行を停止して、前記第Ｉ開放メッセージの発行後に、前記第１プロセッサノード（２０－ｉ）に宛てて書き込み要求メッセージを発行するステップと、前記第Ｉ開放メッセージを受けた前記第１プロセッサノード（２０－ｉ）の前記メモリコントローラ（４０－ｉ）は、前記リクエストロック状態情報に代えてライトロック状態情報を前記第１プロセッサノード（１－ｉ）の前記ディレクトリ（２０－ｉ）に格納し、第Ｉ書き込み許可メッセージを前記第１入出力コントローラに宛てて出力するステップとを含んでいる。

【発明の効果】

【００３４】

以上の説明により、本発明のマルチプロセッサシステム及びメモリアクセス処理方法によれば、入出力コントローラが、異なるプロセッサノードを宛先とする入出力デバイスからの複数のライトメッセージを連続して処理することができる。

本発明のマルチプロセッサシステム及びメモリアクセス処理方法によれば、入出力コントローラが複数のライトメッセージを連続して処理するため、入出力コントローラがライトメッセージの処理を行う場合に要する時間を従来のそれよりも短くすることができる。

【発明を実施するための最良の形態】

【００３５】

以下に添付図面を参照して、本発明のマルチプロセッサシステムについて詳細に説明する。

【００３６】

図４にマルチプロセッサシステムの構成を示す。本発明のマルチプロセッサシステムは、複数のプロセッサノード１－１～１－ｍ（ｍは１以上の整数）と、複数の入出力ノード

３－１～３－ｎ（ｎは１以上の整数）とを具備している。複数のプロセッサノード１－１～１－ｍと、複数の入出力ノード３－１～３－ｎとは、ネットワーク２に接続され、外部からのクロックに応じて動作する。プロセッサノード１－ｉ（ｉ＝１、２、…、ｍ）には、複数のプロセッサ１０－ｉ－１、１０－ｉ－２と、ディレクトリ２０－ｉと、主記憶部（メモリ）３０－ｉと、メモリコントローラ４０－ｉとが設けられている。メモリコントローラ４０－ｉは、プロセッサ１０－ｉ－１、１０－ｉ－２とディレクトリ２０－ｉと主記憶部３０－ｉとに接続されている。入出力ノード３－ｊ（ｊ＝１、２、…、ｎ）には、入出力コントローラ５０－ｊと、外部からの命令によりメッセージを発行する複数の入出力デバイス６０－ｊ－１、６０－ｊ－２とが設けられている。メッセージは、コマンドの種類を表すコマンド種別と、アドレスとを含み、例えば、メッセージがライトメッセージである場合、コマンド種別はライトを表す。また、ネットワーク２はメッセージの配送を行い、２点間のメッセージの順序を保障する。

【００３７】

主記憶部３０－ｉには、複数のデータが格納されている。複数のデータの各々は、その内容を表す値を含んでいる。

ディレクトリ２０－ｉは、主記憶部３０－ｉに格納されている各データの一貫性制御に関する情報を、例えば１２８バイトのブロック単位、で保持している。各ブロックの情報としては、一貫性制御に関する状態情報を含む。この状態情報には、アクセス要求を受け付けることができるフリー状態情報と、他のアクセス要求を受け付けることができないライトロック状態情報が含まれる。

ここで、一貫性制御（一貫性処理）について説明する。マルチプロセッサシステムでは、複数のプロセッサが存在する。また、複数のプロセッサがそれぞれキャッシュを具備し、データのコピーを保持する。そのため、ひとつのデータについて、メモリの値と、コピーを取った複数のキャッシュ値とを一致させる処理が必要になる。この一致させる処理としては、コピーを無効化する処理が例示される。このように、データの値を一致させる、即ち、データの一貫性を取ることを一貫性制御（一貫性処理）と呼ぶ。

データの一貫性制御に関する情報については、後述の実施例にて説明する。

【００３８】

本発明は、入出力コントローラ５０－ｊとメモリコントローラ４０－ｉの間での一連のメッセージを工夫することで、ホームを異にするライトメッセージの連続発行を可能とするものである。

いま、入出力ノード３－１～３－ｎのうちの第１入出力ノード（入出力ノード３－１とする）の入出力デバイス６０－１－１、６０－１－２）によって１番目からＭ番目（Ｍは１以上の整数）までのＭ個のデータに対するＭ個のライトメッセージが発行されたものとする。このとき、複数の入出力コントローラのうちの第１入出力コントローラ（入出力ノード３－１の入出力コントローラ５０－１）がこれらＭ個のライトメッセージを受け取ると、前記Ｍ個のそれぞれのデータに対するＭ個のライトトランザクションを開始する。

Ｍ個のデータのうちの第Ｉデータ（Ｉ＝１、２、…、Ｍ）は、プロセッサノード１－１～１－ｍのうちの１つのプロセッサノード（プロセッサノード１－ｉとする）をホームとするデータである。Ｍ個のライトメッセージのうちの第Ｉライトメッセージは、プロセッサノード１－ｉの主記憶部３０－ｉに格納された複数のデータのうちの第Ｉデータの値を第Ｉライトメッセージで指定される値に更新するための命令である。以降、第Ｉライトメッセージを例にとり説明する。

【００３９】

第Ｉライトメッセージを受けた入出力コントローラ５０－１は、第Ｉライトトランザクションを開始して、第Ｉ書き込み要求メッセージを、ネットワーク２を介してプロセッサノード１－ｉに出力する。プロセッサノード１－ｉのメモリコントローラ４０－ｉは、第Ｉ書き込み要求メッセージを受け取ったとき、第Ｉデータに対して、フリー状態情報に代えて、第Ｉデータに対する他の書き込み要求メッセージを受け付けることができないライトロック状態情報をプロセッサノード１－ｉのディレクトリ２０－ｉに格納し、第Ｉ書き

込み要求メッセージに対して第Ｉ書き込み許可メッセージを入出力ノード３－１にネットワーク２を介して出力する。

【００４０】

入出力ノード３－１の入出力コントローラ５０－１は、第Ｉ書き込み許可メッセージを受け取ったとき、図１４に示されるような発行処理を実行する。

【００４１】

まず、入出力コントローラ５０－１は、第１から第Ｉ書き込み許可メッセージまでのＩ個の書き込み許可メッセージを既に受け取っているか否かを検査する（図１４のステップＳ１）。ここで、入出力コントローラ５０－１は、上記のＩ個の書き込み許可メッセージを未だ受け取っていない場合（図１４のステップＳ１－ＮＯ）、発行処理を終了し、次の書き込み許可メッセージの到着を待つ。入出力コントローラ５０－１は、上記のＩ個の書き込み許可メッセージを既に受け取っている場合（図１４のステップＳ１－ＹＥＳ）、第Ｉライトメッセージで指定される値を含む第Ｉ更新メッセージをプロセッサノード１－ｉにネットワーク２を介して出力する（図１４のステップＳ２）。

【００４２】

次に、入出力コントローラ５０－１は、Ｉ＝Ｉ＋１とし、ＩがＭ以下であるかどうかを検査する（図１４のステップＳ３、Ｓ４）。ＩがＭ以下である、即ち、第Ｉライトメッセージを受け取って第Ｉライトトランザクションを開始していれば（図１４のステップＳ４－ＹＥＳ）、ステップＳ１を実行する。ＩがＭより大きい、即ち、まだ第Ｉライトメッセージを受け取っていなければ（図１４のステップＳ４－ＮＯ）処理を終了する。

【００４３】

このように、入出力コントローラ５０－１は、ステップＳ１～Ｓ４を繰り返し実行し、当該ライトトランザクションおよび先行するライトトランザクションが全て書き込み許可メッセージを受け取っていれば、更新メッセージを発行する。

【００４４】

プロセッサノード１－ｉのメモリコントローラ４０－ｉは、第Ｉ更新メッセージを受け取ったとき、第Ｉデータに対してライトロック状態情報に代えてフリー状態情報をプロセッサノード１－ｉのディレクトリ２０－ｉに格納すると共に、プロセッサノード１－ｉの主記憶部３０－ｉに格納された第Ｉデータの値を第Ｉ更新メッセージで指定される値に更新する。

【００４５】

このように、本発明のマルチプロセッサシステムによれば、入出力コントローラ５０－ｊが、異なるプロセッサノード１－ｉを宛先とする入出力デバイス６０－ｊ－１、６０－ｊ－２からの複数のライトメッセージを連続して処理することができる。

【００４６】

図６を参照しながら、順序制約のあるライトメッセージが連続して発行された場合の動作を具体的に説明する。図は、入出力デバイス６０－１－１が、Ｍ個（Ｍ＝３）のデータとしてデータＡ、Ｂ、Ｃに対するライトメッセージをそれぞれステップ１、２、３で発行した場合の動作を示している。ここで、データＡおよびデータＢはプロセッサノード１－１をホームとし、データＣはプロセッサノード１－２をホームとするデータであるとする。また、１ステップは１クロックに対応する。

【００４７】

入出力コントローラ５０－１は、ステップ２にて、第１ライトメッセージとしてライトＡメッセージを受け取る。このとき、入出力コントローラ５０－１は、ライトＡトランザクションを開始し、ステップ３にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に、第１書き込み要求メッセージとして書き込み要求Ａメッセージを出力する。

メモリコントローラ４０－１は、ステップ４にて、入出力コントローラ５０－１からの書き込み要求Ａメッセージを受け取ったとき、ディレクトリ２０－１が保持するデータＡの状態情報をフリー状態情報からライトロック状態情報に更新する。また、メモリコント

ローラ４０－１は、ステップ５にて、入出力コントローラ５０－１に宛ててネットワーク２に、第１書き込み許可メッセージとして書き込み許可Ａメッセージを出力する。

入出力コントローラ５０－１は、ステップ６にて、書き込み許可Ａメッセージを受け取る。このとき、先行するライトメッセージが存在しないので、入出力コントローラ５０－１は、ステップ７にて、第１ライトメッセージで指定される値を含む第１更新メッセージとして、更新Ａメッセージをメモリコントローラ４０－１に宛ててネットワーク２に出力する。

メモリコントローラ４０－１は、ステップ８にて、入出力コントローラ５０－１からの更新Ａメッセージを受け取る。このとき、メモリコントローラ４０－１は、ディレクトリ２０－１が保持するデータＡの状態情報をライトロック状態情報からフリー状態情報に更新し、主記憶部３０－１のデータの値を更新Ａメッセージで指定される値に更新する（入出力コントローラ５０－１からのライトＡメッセージであるデータＡを格納する）。

【００４８】

入出力コントローラ５０－１は、ステップ３にて、第２ライトメッセージとしてライトＢメッセージを受け取る。このとき、入出力コントローラ５０－１は、ライトＢトランザクションを開始し、ステップ４にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に、第２書き込み要求メッセージとして書き込み要求Ｂメッセージを出力する。

メモリコントローラ４０－１は、ステップ５にて、入出力コントローラ５０－１からの書き込み要求Ｂメッセージを受け取ったとき、ディレクトリ２０－１が保持するデータＢの状態情報をフリー状態情報からライトロック状態情報に更新する。また、メモリコントローラ４０－１は、ステップ６にて、入出力コントローラ５０－１に宛ててネットワーク２に、第２書き込み許可メッセージとして書き込み許可Ｂメッセージを出力する。

入出力コントローラ５０－１は、ステップ７にて、書き込み許可Ｂメッセージを受け取る。このとき、入出力コントローラ５０－１は、先行するライトＡトランザクションの進捗を検査する。入出力コントローラ５０－１は、ステップ７より前のステップ６で既に第１書き込み許可メッセージである書き込み許可Ａメッセージを受け取っている。よって、入出力コントローラ５０－１は、ステップ８にて、第２ライトメッセージで指定される値を含む第２更新メッセージとして、更新Ｂメッセージをメモリコントローラ４０－１に宛ててネットワーク２に出力する。

メモリコントローラ４０－１は、ステップ９にて、入出力コントローラ５０－１からの更新Ｂメッセージを受け取る。このとき、メモリコントローラ４０－１は、ディレクトリ２０－１が保持するデータＢの状態情報をライトロック状態情報からフリー状態情報に更新し、主記憶部３０－１のデータの値を更新Ｂメッセージで指定される値に更新する（入出力コントローラ５０－１からのライトＢメッセージであるデータＢを格納する）。

【００４９】

入出力コントローラ５０－１は、ステップ４にて、第３ライトメッセージとしてライトＣメッセージを受け取る。このとき、入出力コントローラ５０－１は、ライトＣトランザクションを開始し、ステップ５にて、ホームのメモリコントローラ４０－２に宛ててネットワーク２に、第３書き込み要求メッセージとして書き込み要求Ｃメッセージを出力する。

メモリコントローラ４０－２は、ステップ６にて、入出力コントローラ５０－１からの書き込み要求Ｃメッセージを受け取ったとき、ディレクトリ２０－２が保持するデータＣの状態情報をフリー状態情報からライトロック状態情報に更新する。また、メモリコントローラ４０－１は、ステップ７にて、入出力コントローラ５０－１に宛ててネットワーク２に、第３書き込み許可メッセージとして書き込み許可Ｃメッセージを出力する。

入出力コントローラ５０－１は、ステップ８にて、書き込み許可Ｃメッセージを受け取る。このとき、入出力コントローラ５０－１は、先行するライトＡトランザクションおよびライトＢトランザクションの進捗を検査する。入出力コントローラ５０－１は、ステップ８より前のステップ６で既に書き込み許可Ａメッセージを受け取り、ステップ８より前

のステップ７で既に書き込み許可Ｂメッセージを受け取っている。よって、入出力コントローラ５０－１は、ステップ９にて、第３ライトメッセージで指定される値を含む第３更新メッセージとして、更新Ｃメッセージをメモリコントローラ４０－２に宛ててネットワーク２に出力する。

メモリコントローラ４０－２は、ステップ１０にて、入出力コントローラ５０－１からの更新Ｃメッセージを受け取る。このとき、メモリコントローラ４０－２は、ディレクトリ２０－２が保持するデータＣの状態情報をライトロック状態情報からフリー状態情報に更新し、主記憶部３０－２のデータの値を更新Ｃメッセージで指定される値に更新する（入出力コントローラ５０－１からのライトＣメッセージであるデータＣを格納する）。

【００５０】

入出力コントローラ５０－ｊには、図５に示されるように、セレクタ７１－ｊと、メッセージ格納キュー７２－ｊと、ライトポインタ７３－ｊと、リードポインタＡ７４－ｊと、リードポインタＢ７５－ｊとが設けられている。入出力コントローラ５０－ｊは、複数の入出力デバイス６０－ｊ－１、６０－ｊ－２からのメッセージをセレクタ７１－ｊで調停してメッセージ格納キュー７２－ｊに書き込む。メッセージ格納キュー７２－ｊにおける書き込みの制御はライトポインタ７３－ｊを用いて行われ、読み出しの制御はリードポインタＡ７４－ｊとリードポインタＢ７５－ｊの二つを用いて行われる。また、入出力コントローラ５０－ｊは、ネットワーク２から送られてくるメッセージに応じて許可フラグ７６－１をメッセージ格納キュー７２－ｊに設定（格納）し、メッセージ格納キュー７２－１の制御に該許可フラグ７６－１も用いられる。

ここで、上記図６に示した動作フローでの入出力コントローラ５０－１のメッセージ格納キュー７２－１、ライトポインタ７３－１、リードポインタＡ７４－１、リードポインタＢ７５－１、許可フラグ７６－１の値の遷移を図７Ａ～図７Ｉに示す。

【００５１】

図７Ａに示されるように、初期状態として、ステップ１にて、ライトポインタ７３－１、リードポインタＡ７４－１、リードポインタＢ７５－１は全て「０」を指している。

【００５２】

入出力コントローラ５０－１は、ステップ２にて、入出力デバイス６０－１－１からライトＡメッセージを受け取ると、図７Ｂに示されるように、メッセージ格納キュー７２－１のライトポインタ７３－１が指すエントリ「０」にライトＡメッセージを書き込み、同じく許可フラグ７６－１のエントリ「０」の値を「０」に設定し、ライトポインタ７３－１を「１」に更新する。

【００５３】

入出力コントローラ５０－１は、ステップ３にて、入出力デバイス６０－１－１からライトＢメッセージを受け取ると、図７Ｃに示されるように、メッセージ格納キュー７２－１のライトポインタ７３－１が指すエントリ「１」にライトＢメッセージを書き込み、同じく許可フラグ７６－１のエントリ「１」の値を「０」に設定し、ライトポインタ７３－１を「２」に更新する。

また、ステップ２にてメッセージ格納キュー７２－１に有効なエントリが存在することになる。このため、図７Ｃに示されるように、入出力コントローラ５０－１は、ステップ３にて、リードポインタＡ７４－１が示すエントリ「０」の情報により、書き込み要求Ａメッセージをネットワーク２に出力し、リードポインタＡ７４－１の値を「１」に更新する。

【００５４】

入出力コントローラ５０－１は、ステップ４にて、入出力デバイス６０－１－１からライトＣメッセージを受け取る。このとき、図７Ｄに示されるように、入出力コントローラ５０－１は、メッセージ格納キュー７２－１のライトポインタ７３－１が指すエントリ「２」にライトＣメッセージを書き込み、同じく許可フラグ７６－１のエントリ「２」の値を「０」に設定し、ライトポインタ７３－１を「３」に更新する。

また、ステップ３にてメッセージ格納キュー７２－１に有効なエントリが存在すること

になる。このため、図７Ｄに示されるように、入出力コントローラ５０－１は、ステップ４にて、リードポインタＡ７４－１が示すエントリ「１」の情報により、書き込み要求Ｂメッセージをネットワーク２に出力し、リードポインタＡ７４－１の値を「１」から「２」に更新する。

【００５５】

ステップ４にてメッセージ格納キュー７２－１に有効なエントリが存在することになる。このため、図７Ｅに示されるように、入出力コントローラ５０－１は、ステップ５にて、リードポインタＡ７４－１が示すエントリ「２」の情報により、書き込み要求Ｃメッセージをネットワーク２に出力し、リードポインタＡ７４－１の値を「２」から「３」に更新する。

【００５６】

入出力コントローラ５０－１は、ステップ６にて、書き込み許可Ａメッセージを受け取ったとき、図７Ｆに示されるように、許可フラグ７６－１のデータＡに該当するエントリ「０」の値を「０」から「１」に更新する。そしてリードポインタＢ７５－１が指すエントリと書き込み許可メッセージを受け取ったデータＡが格納されているエントリが一致しているかどうかを検査する。この値が一致するということは、先行するライトトランザクションが存在しない、あるいは既に書き込み許可メッセージを受け取って更新メッセージを発行し完了していることを示している。ここでは「０」で一致するので、更新メッセージの発行処理を行う。次の書き込み許可メッセージの到着を待つ。

【００５７】

入出力コントローラ５０－１は、ステップ６にてリードポインタＢ７５－１が指すメッセージ格納キュー７２－１のエントリ「０」の値を読み出し、ステップ７にて、ネットワーク２に更新Ａメッセージを出力する。このとき、図７Ｇに示されるように、入出力コントローラ５０－１は、リードポインタＢ７５－１の値を「０」から「１」に更新する。

次に、リードポインタＢ７５－１とライトポインタ７３－１の値を比較し、リードポインタＢ７５－１の値が小さく未完了の書き込みが存在することを示す場合、未完了のライトトランザクションで既に書き込み許可メッセージを受け取り済みのものが存在するかどうかを検査する。ここでは、リードポインタＢ７５－１の値「１」はライトポインタ７３－１の値「３」より小さいので、リードポインタＢ７５－１が指す許可フラグ７６－１のエントリ「１」の値を読み出す。値は「０」であり書き込み許可メッセージをまだ受け取っていないことを示すので、エントリ「１」の更新メッセージ発行処理は行わず次の書き込み許可メッセージの到着を待つ。

【００５８】

入出力コントローラ５０－１は、ステップ７にて、許可Ｂメッセージを受け取ると、図７Ｇに示されるように、許可フラグ７６－１のデータＢに該当するエントリ「１」の値を「０」から「１」に更新する。そしてリードポインタＢ７５－１が指すエントリと書き込み許可メッセージを受け取ったデータＢが格納されているエントリが一致しているかどうかを検査する。ここでは「１」で一致するので、更新メッセージの発行処理を行う。

【００５９】

入出力コントローラ５０－１は、ステップ７にてリードポインタＢ７５－１が指すメッセージ格納キュー７２－１のエントリ「１」の値を読み出し、ステップ８にて、ネットワーク２にライトＢメッセージを出力する。このとき、図７Ｈに示されるように、入出力コントローラ５０－１は、リードポインタＢ７５－１の値を「１」から「２」に更新する。

次に、リードポインタＢ７５－１の値「２」とライトポインタ７３－１の値「３」を比較すると、リードポインタＢ７５－１の値が小さいので、リードポインタＢ７５－１が指す許可フラグ７６－１のエントリ「２」の値を読み出す。値は「０」であり書き込み許可メッセージをまだ受け取っていないことを示すので、エントリ「２」の更新メッセージ発行処理は行わず次の書き込み許可メッセージの到着を待つ。

【００６０】

入出力コントローラ５０－１は、ステップ８にて、書き込み許可Ｃメッセージを受け取

ると、図７Ｈに示されるように、許可フラグ７６－１のデータＣに該当するエントリ「２」の値を「０」から「１」に更新する。そしてリードポインタＢ７５－１が指すエントリと書き込み許可メッセージを受け取ったデータＣが格納されているエントリが一致しているかどうかを検査する。ここでは「２」で一致するので、更新メッセージの発行処理を行う。

【００６１】

入出力コントローラ５０－１は、ステップ８にてリードポインタＢ７５－１が指すメッセージ格納キュー７２－１のエントリ「２」の値を読み出し、ステップ９にて、ネットワーク２に更新Ｃメッセージを出力する。このとき、図７Ｉに示されるように、入出力コントローラ５０－１は、リードポインタＢ７５－１の値を「２」から「３」に更新する。

次に、リードポインタＢ７５－１の値「３」とライトポインタ７３－１の値「３」を比較すると、値が一致し未完了のライトトランザクションが存在しないことを示すので、処理を終了する。

【００６２】

上記のように動作することで、ホームを同じにするライトトランザクションの順序も、異なるホームへのライトトランザクションの順序も保障することができる。

ライトＢトランザクションとライトＡトランザクションの順序関係は次のような理由で保障される。更新Ｂメッセージを発行する時点で、既に書き込み許可Ａメッセージを受け取り済みであり、ホームのディレクトリ２０－１の状態情報が、ライトロック状態に遷移していることが保障される。ライトロック状態に遷移していれば、プロセッサや入出力デバイスは更新Ａメッセージを受け取ってデータＡの値が更新されフリー状態に遷移した後の更新された値しか読み出すことができない。よって、データＢの更新された値が読み出せる時点で、データＡの更新された値しか読み出すことができないので、順序が保障されたことになる。

【００６３】

ライトＣトランザクションとライトＢトランザクションの順序関係も同じな理由で保障される。更新Ｃメッセージを発行する時点で、既に書き込み許可Ｂメッセージを受け取り済みであり、ホームのディレクトリ２０－１の状態情報が、ライトロック状態に遷移していることが保障される。ライトロック状態に遷移していれば、プロセッサや入出力デバイスは更新Ｂメッセージを受け取ってデータＢの値が更新されフリー状態に遷移した後の更新された値しか読み出すことができない。よって、データＣの更新された値が読み出せる時点で、データＢの更新された値しか読み出すことができないので、順序が保障されたことになる。

【００６４】

また、上記処理は１０ステップで完了しており、従来技術と比べて入出力デバイス６０－ｊ－１からのライトメッセージの処理性能を向上させることができている。

【００６５】

図８に示すように、上記の入出力コントローラ５０－ｊは、入出力ノード３－ｊ内に設けられているが、入出力コントローラ５０－ｊとして入出力コントローラ５２－ｊがネットワーク２に接続されていてもよい。この場合、入出力ノード３－ｊには、上記の入出力セレクタ７１－ｊに対応する入出力セレクタ５１－ｊが設けられていることが好ましい。入出力セレクタ５１－ｊは、複数の入出力デバイス６０－ｊ－１、６０－ｊ－２が出力するメッセージを調停してネットワーク２に出力し、入出力コントローラ５２－ｊは、ネットワーク２に出力されたメッセージを受けて、図４に示した入出力コントローラ５０－ｊと同様の処理を行う構成でも良い。

【００６６】

また、上記図８に示した入出力コントローラ５２－ｊは、ネットワーク２に接続されるのではなく、入出力コントローラ５２－ｉ（図示しない）としてプロセッサノード１－ｉ内に設けられていても良い。この場合、入出力ノード３－ｊには、上記の入出力セレクタ５１－ｊが設けられていることが好ましい。

【００６７】
また、図４や図８に示す構成で、ｍとｎとが等しく、プロセッサノード１－ｉと入出力ノード３－ｊとで一つのノード（図示しない）を構成してもよい。この場合、そのノードには、ネットワーク２が接続され、上記の入出力セレクタ５１－ｊが設けられていることが好ましい。
【００６８】
以上の説明により、本発明のマルチプロセッサシステムによれば、入出力コントローラ５０－ｊが、異なるプロセッサノード１－ｉを宛先とする入出力デバイス６０－ｊ－１、６０－ｊ－２からの複数のライトメッセージを連続して処理することができる。
本発明のマルチプロセッサシステムによれば、入出力コントローラ５０－ｊが複数のライトメッセージを連続して処理するため、入出力コントローラ５０－ｊがライトメッセージの処理を行う場合に要する時間を従来のそれよりも短くすることができる。
【００６９】
以降、入出力コントローラ５０－ｊとメモリコントローラ４０－ｉ間のやりとりをより詳細に説明する。
【実施例１】
【００７０】
上述のように、ディレクトリ２０－ｉは、主記憶部３０－ｉに格納されているデータの一貫性制御に関する情報を、例えば１２８バイトのブロック単位、で保持している。データの一貫性制御に関する情報は、主記憶部３０－ｉの各ブロックの状態情報（データの一貫性制御に関する状態情報）と、マップ情報とを含んでいる。
【００７１】
格納されるブロックの状態情報（データの一貫性制御に関する状態情報）には、上述のように、他のアクセス要求を受け付けることができるフリー状態情報と、他のアクセス要求を受け付けることができないライトロック状態情報が含まれる。
フリー状態情報は、例えば、Ｕｎｃａｃｈｅｄ、Ｃｌｅａｎ、Ｄｉｒｔｙ（以降Ｕ、Ｃ、Ｄとも略す）の３つの状態情報からなる。
Ｕｎｃａｃｈｅｄは、複数のプロセッサノード１－１～１－ｍのうちのどのプロセッサノードもデータをキャッシングしていないことを示す。
Ｃｌｅａｎは、複数のプロセッサノード１－１～１－ｍのうちの少なくとも１つのプロセッサノードがデータをキャッシングしていることを示す。
Ｄｉｒｔｙは、複数のプロセッサノード１－１～１－ｍのうちのある１つのプロセッサノードがデータをキャッシングし、最新のデータはその１つのプロセッサノードにのみ存在することを示す。
【００７２】
また、データの一貫性制御に関する状態情報には、他のアクセス要求を受け付けることができない状態情報として、要求（リクエスト）されたメッセージのみを受け付けることができるリクエストロック状態情報が含まれるものとする。本発明において、ライトロック状態情報は、このリクエストロック状態情報と同一にしてもよいし異なる二つの状態情報としても良い。以降の説明では二つの異なる状態情報（以降リクエストロック状態情報をＲ、ライトロック状態情報をＷとも略す）として存在する場合を例にとり説明する。
【００７３】
また、マップ情報は、各プロセッサノード１－ｉが該ブロックのデータをキャッシングしているかどうかを示す情報である。プロセッサノード１数分のビットを用いて、マップ情報について説明する。図４の構成例としてｍを３とする。即ち、プロセッサノードが３ノード存在するので（プロセッサノード１－１、１－２、１－３）、３ビットで表現するものとする。
例えば“０００”はどのプロセッサノード１－１、１－２、１－３もデータをキャッシングしていないことを示す。“００１”はプロセッサノード１－１がデータをキャッシングしていることを示す。“０１０”はプロセッサノード１－２がデータをキャッシングし

ていることを示す。“１００”はプロセッサノード１－３がデータをキャッシングしていることを示す。同様に、“１１０”はプロセッサノード１－２とプロセッサノード１－３がデータをキャッシングしていることを示す。

【００７４】

図９を参照しながら、本発明の第１実施例に係るマルチプロセッサシステムの動作として、メモリコントローラ４０－１が書き込み要求Ａメッセージを受けたときに、ディレクトリ２０－１の該当するブロックの状態情報がＵｎｃａｃｈｅｄ状態情報であった場合の動作について説明する。図は、入出力デバイス６０－１－１が、データＡに対するライトメッセージをステップ１で発行した場合の動作を示している。ここで、データＡはプロセッサノード１－１をホームとするデータであるとする。また、１ステップは１クロックに対応する。

【００７５】

入出力コントローラ５０－１は、ステップ２にて、ライトＡメッセージを受け取る。このとき、入出力コントローラ５０－１は、ステップ３にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に書き込み要求Ａメッセージを出力する。

メモリコントローラ４０－１は、ステップ４にて、入出力コントローラ５０－１からの書き込み要求Ａメッセージを受け取り、ディレクトリ２０－１が保持するデータＡの状態情報をフリー状態情報からライトロック状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｕ、０００”から“Ｗ、０００”に更新する。ここで、“Ｕ、０００”は状態情報がＵであることを示し、マップ情報が“０００”であることを示す。また、メモリコントローラ４０－１は、ステップ５にて、入出力コントローラ５０－１に宛ててネットワーク２に書き込み許可Ａメッセージを出力する。

入出力コントローラ５０－１は、ステップ６にて、書き込み許可Ａメッセージを受け取る。次に、ステップ７にて、更新Ａメッセージをメモリコントローラ４０－１に宛ててネットワーク２に出力する。

メモリコントローラ４０－１は、ステップ８にて、入出力コントローラ５０－１からの更新Ａメッセージを受け取り、主記憶部３０－１のデータの値を更新Ａメッセージで指定される値に更新する。そして、ディレクトリ２０－１が保持するデータＡの状態情報をライトロック状態情報からフリー状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｗ、０００”から“Ｕ、０００”に更新する。

【００７６】

図１０を参照しながら、本発明の第１実施例に係るマルチプロセッサシステムの動作として、メモリコントローラ４０－１が書き込み要求Ａメッセージを受けたときに、ディレクトリ２０－１の該当するブロックの状態情報がＣｌｅａｎ状態情報で、プロセッサノード１－２とプロセッサノード１－３がデータをキャッシングしている場合の動作について説明する。図は、入出力デバイス６０－１－１が、データＡに対するライトメッセージをステップ１で発行した場合の動作を示している。ここで、データＡはプロセッサノード１－１をホームとするデータであるとする。また、１ステップは１クロックに対応する。

【００７７】

入出力コントローラ５０－１は、ステップ２にて、ライトＡメッセージを受け取る。このとき、入出力コントローラ５０－１は、ステップ３にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に書き込み要求Ａメッセージを出力する。

メモリコントローラ４０－１は、ステップ４にて、入出力コントローラ５０－１からの書き込み要求Ａメッセージを受け取り、ディレクトリ２０－１が保持するデータＡの状態情報をフリー状態情報からライトロック状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｃ、１１０”から“Ｗ、０００”に更新する。ここで、“Ｃ、１１０”は状態情報がＣであることを示し、マップ情報が“１１０”であることを示す。また、メモリコントローラ４０－１は、ステップ５にて、入出力コントローラ５０－１に宛ててネットワーク２に応答Ａメッセージを出力し、メモリコントローラ４０－２とメモリコントローラ４０－３に宛てて無効化Ａメッセージを出力する。このとき、応答Ａメッセ

ージには、キャッシングしているプロセッサノードの数（この例では２）が付加される。
入出力コントローラ５０－１は、ステップ６にて、応答Ａメッセージを受け取る。また、メモリコントローラ４０－２、メモリコントローラ４０－３は、ステップ６にて、それぞれ無効化Ａメッセージを受け取り、それぞれ当該プロセッサノード１－２、プロセッサノード１－３でキャッシングしているデータＡを無効化する。そして、メモリコントローラ４０－２、メモリコントローラ４０－３は、ステップ７にて、それぞれネットワーク２に入出力コントローラ５０－１に宛てて無効化完了Ａメッセージを出力する。
入出力コントローラ５０－１は、ステップ８にて、応答Ａメッセージに付加されている数の無効化完了Ａメッセージを受け取った時点で、書き込み許可Ａメッセージを受け取ったと解釈（認識）する。そして、ステップ９にて、更新Ａメッセージをメモリコントローラ４０－１に宛ててネットワーク２に出力する。
メモリコントローラ４０－１は、ステップ１０にて、入出力コントローラ５０－１からのライトＡメッセージを受け取り、主記憶部３０－１のデータの値を更新Ａメッセージで指定される値に更新する。そして、ディレクトリ２０－１が保持するデータＡの状態情報をライトロック状態情報からフリー状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｗ、０００”から“Ｕ、０００”に更新する。

【００７８】

図１１を参照しながら、本発明の第１実施例に係るマルチプロセッサシステムの動作として、メモリコントローラ４０－１が書き込み要求Ａメッセージを受けたときに、ディレクトリ２０－１の該当するブロックの状態情報がＤｉｒｔｙ状態情報で、プロセッサノード１－２がデータをキャッシングしている場合の動作について説明する。図は、入出力デバイス６０－１－１が、データＡに対するライトメッセージをステップ１で発行した場合の動作を示している。ここで、データＡはプロセッサノード１－１をホームとするデータであるとする。また、１ステップは１クロックに対応する。

【００７９】

入出力コントローラ５０－１は、ステップ２にて、ライトＡメッセージを受け取る。このとき、入出力コントローラ５０－１は、ステップ３にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に書き込み要求Ａメッセージを出力する。
メモリコントローラ４０－１は、ステップ４にて、入出力コントローラ５０－１からの書き込み要求Ａメッセージを受け取り、ディレクトリ２０－１が保持するデータＡの状態情報をフリー状態情報からリクエストロック状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｄ、０１０”から“Ｒ、０１０”に更新する。ここで、“Ｒ、０１０”は状態情報がＲであることを示し、マップ情報が“０１０”であることを示す。また、メモリコントローラ４０－１は、ステップ５にて、入出力コントローラ５０－１に宛ててネットワーク２に書き戻し要求Ａメッセージを出力する。ここで、リクエストロック状態情報である“Ｒ，０１０”への更新ではなく、ライトロック状態情報である“Ｗ，０００”への更新であっても構わない。
メモリコントローラ４０－２は、ステップ６にて、メモリコントローラ４０－１からの書き戻し要求Ａメッセージを受け取り、該プロセッサノード１－２でキャッシングしているデータＡの書き戻しを行い、ステップ７にて、メモリコントローラ４０－１に宛ててネットワーク２に書き戻しＡメッセージを出力する。
メモリコントローラ４０－１は、ステップ８にて、リクエストロック状態情報によりリクエストされたメッセージとして、書き戻しＡメッセージをメモリコントローラ４０－２から受け取り、ディレクトリ２０－１が保持するデータＡの状態情報をリクエストロック状態情報からライトロック状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｒ、０１０”から“Ｗ、０００”に更新する。また、メモリコントローラ４０－１は、ステップ９にて、入出力コントローラ５０－１に宛ててネットワーク２に書き込み許可Ａメッセージを出力する。
入出力コントローラ５０－１は、ステップ１０にて、書き込み許可Ａメッセージを受け取る。次に、ステップ１１にて、更新Ａメッセージをメモリコントローラ４０－１に宛て

てネットワーク２に出力する。
メモリコントローラ４０－１は、ステップ１１にて、入出力コントローラ５０－１からの更新Ａメッセージを受け取り、主記憶部３０－１のデータの値を更新Ａメッセージで指定される値に更新する。そして、ディレクトリ２０－１が保持するデータＡの状態情報をライトロック状態情報からフリー状態情報に更新する。即ち、ディレクトリ２０－１の該ブロックの値を“Ｗ、０００”から“Ｕ、０００”に更新する。
【００８０】
図１２を参照しながら、本発明の第１実施例に係るマルチプロセッサシステムの動作として、メモリコントローラ４０－１が書き込み要求Ａメッセージを受けたときに、ディレクトリ２０－１の該当するブロックの状態情報がリクエストロック状態情報あるいはライトロック状態情報であった場合の動作について説明する。図は、入出力デバイス６０－１－１が、データＡに対するライトメッセージをステップ１で発行した場合の動作を示している。ここで、データＡはプロセッサノード１－１をホームとするデータであるとする。また、１ステップは１クロックに対応する。
【００８１】
入出力コントローラ５０－１は、ステップ２にて、ライトＡメッセージを受け取る。このとき、入出力コントローラ５０－１は、ステップ３にて、ホームのメモリコントローラ４０－１に宛ててネットワーク２に書き込み要求Ａメッセージを出力する。
メモリコントローラ４０－１は、ステップ４にて、入出力コントローラ５０－１からの書き込み要求Ａメッセージを受け取る。このとき、ディレクトリ２０－１が保持するデータＡの状態情報は、リクエストロック状態情報“Ｒ”あるいはライトロック状態情報“Ｗ”である。このため、メモリコントローラ４０－１は、ステップ５にて、入出力コントローラ５０－１に宛ててネットワーク２に不許可Ａメッセージを出力する。
入出力コントローラ５０－１は、ステップ６にて、不許可Ａメッセージを受け取り、ステップ７にて、メモリコントローラ４０－１に宛ててネットワーク２に書き込み要求Ａメッセージを再出力する。ステップ７以降の動作は、これまで図９から図１２を参照しながら説明した動作と同じであるので省略する。
【００８２】
以上のように動作することで、ディレクトリのマップ情報に入出力ノードの分のビットを加えなくて済む。また、ホームが受けたプロセッサからのアクセス要求を、入出力コントローラ５０－ｊに対して転送せずに済み、メモリコントローラや入出力コントローラの構成が複雑にならずに済む。
【実施例２】
【００８３】
本発明のマルチプロセッサシステムの入出力コントローラ５０－ｊは図１３に示すように構成されても良い。この入出力コントローラ５０－ｊは、更に、開放処理ポインタ７８－ｊと開放処理フラグ７９－ｊを有する。初期状態で、開放処理フラグ７９－ｊの値は“０”である。
基本動作は実施例１と同様であるので、ここでは異なる部分のみを説明し、同じ部分は省略する。また、実施の最良の形態と同様に、入出力ノード３－１～３－ｎのうちの第１入出力ノード（入出力ノード３－１とする）の入出力デバイス６０－１－１、６０－１－２）によって１番目からＭ番目（Ｍは１以上の整数）までのＭ個のデータに対するＭ個のライトメッセージが発行されたものとする。以降、第Ｉライトメッセージを例にとり説明する。
【００８４】
メモリコントローラ４０－ｉは、第Ｉデータに対して前記ライトロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、前記第Ｉデータに対する第Ｉ書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、第Ｉ書き込み要求メッセージに対して第Ｉ開放要求メッセージを入出力コントローラ５０－１にネットワーク２を介して出力する。第Ｉ開放要求メッセージを受けた入出力コントローラ５０－１は

、第Ｉ書き込み要求メッセージをメモリコントローラ４０－ｉに宛ててネットワーク２に出力する。また、開放処理フラグ７９－１の値を“０”から“１”に更新し、開放処理ポインタ７８－１の値を、第Ｉ番目のライトトランザクションであることを示す情報“Ｉ”に設定する。そして後述の開放処理を行う。

【００８５】

この開放処理フラグ７９－１は、入出力コントローラが後述の開放処理を行っている状態であるかどうかを示し、開放処理ポインタは、その開放処理がどのライトトランザクションによって引き起こされたものであるかを示す。開放処理ポインタ７８－１の値は、開放処理を引き起こしたライトトランザクションと他のライトトランザクションとの順序関係（先行するものか後続するものか）を判断するのにも用いる。

【００８６】

入出力コントローラ５０－１は、開放処理フラグ７９－１の値が“１”の間、第Ｉライトに後続する第Ｋトランザクション｛Ｋは、Ｋ＝Ｉ＋１、Ｉ＋２、…、Ｍを満たす整数であり、Ｉ＋１は、Ｉ＜（Ｉ＋１）＜Ｍを満たす整数であり、Ｉ＋２は、（Ｉ＋１）＜（Ｉ＋２）＜Ｍを満たす整数である｝に関して、次の処理を行う。

入出力コントローラ５０－１は、未だ第Ｋ書き込み要求メッセージを発行していない場合、第Ｋ書き込み要求メッセージの発行を停止する。

入出力コントローラ５０－１は、既に第Ｋ書き込み要求メッセージを発行し第Ｋ書き込み許可メッセージを受け取っている場合、許可フラグ７６－１の該当するエントリの値を“０”に更新して、第Ｋ開放メッセージを前記第Ｋデータのホームであるプロセッサノード１０―ｋ（ｋ＝１、２、…、ｍ）のメモリコントローラ４０－ｋにネットワーク２を介して出力する。

入出力コントローラ５０－１は、既に第Ｋ書き込み要求メッセージを発行し未だ第Ｋ書き込み許可メッセージを受け取っていない場合、第Ｋ書き込み許可メッセージを受け取ったときに、許可フラグ７６－１の該当するエントリの値を“１”に更新せず、第Ｋ開放メッセージの発行を行う。

第Ｋ開放メッセージを受け取ったメモリコントローラ４０－ｋは、そのディレクトリ２０－ｋが保持するデータの状態情報をライトロック状態情報からフリー状態情報に更新する。即ち、ディレクトリ２０－２の該ブロックの値を“Ｗ、０００”から“Ｕ、０００”に更新する。

既に第Ｋ書き込み要求メッセージを発行したがまだ第Ｋ許可メッセージを受け取っていない場合、受け取るメッセージによって次のように動作する。第Ｋ書き込み許可メッセージを受け取った場合、許可フラグ７６－１の該当するエントリの値は“０”のままとし、第Ｋ開放メッセージをメモリコントローラ４０に発行する。第Ｋ不許可メッセージあるいは第Ｋ開放要求メッセージを受け取った場合はなにもしない。

【００８７】

また、入出力コントローラ５０－１は、開放処理フラグ７９－１の値が“１”の間、第Ｉライトトランザクションに先行する第１～第（Ｉ－１）ライトトランザクションに関しては、実施例１と同じように動作する。ただし、先行する第Ｌライトトランザクション｛Ｌは、１≦Ｌ≦（Ｉ－１）を満たす整数｝で開放要求メッセージを受け取った場合、第Ｌ書き込み要求メッセージをメモリコントローラ４０－１に宛ててネットワーク２に出力する。開放処理ポインタ７８－１の値を第Ｌ番目のライトトランザクションであることを示す情報に設定する。そして、第Ｌライトを基準に、先行するライトトランザクションであるか後続するライトトランザクションであるかを判断し、上記開放処理を行う。

開放処理ポインタ７８－１の値が示すライトトランザクション（ここでは第Ｉライトトランザクションとする）に関して第Ｉ書き込み許可メッセージを受け取ると、第Ｉ更新メッセージをメモリコントローラ４０―ｉに発行する。また、開放処理フラグ７９－１の値を“０”に更新し、開放処理を終了する。

【００８８】

以上のように動作することで、例えば入出力デバイス６０－１－１がライトＡメッセー

ジ、ライトＣメッセージの順に発行し、別の入出力デバイス６０－２－２がライトＣメッセージ、ライトＡメッセージの順に発行した場合のデッドロックの危険性を回避することができる。

【実施例３】

【００８９】

本発明のマルチプロセッサシステムは、以下に示すような動作を実行することもできる。

ここで、ディレクトリ２０－ｉに格納される状態情報に、さらにある特定のメッセージの処理のみを受け付ける割り込み可ライトロック状態情報（以降Ｗｉと記す）が存在する。また、割り込み可ライトロック状態時にどの入出力ノード３－ｊが発行したメッセージによって遷移したかを示すフィールドが加わる。以降例えば“Ｗｉ、０００、２”とした場合、入出力ノード３－２が発行したメッセージによってＷｉ状態に遷移したことを示す。

基本動作は実施例１と同様であるので、ここでは異なる部分のみを説明し、同じ部分は省略する。また、実施の最良の形態と同様に、入出力ノード３－１～３－ｎのうちの第１入出力ノード（入出力ノード３－１とする）の入出力デバイス６０－１－１、６０－１－２）によって１番目からＭ番目（Ｍは１以上の整数）までのＭ個のデータに対するＭ個のライトメッセージが発行されたものとする。以降、第Ｉライトメッセージを例にとり説明する。

【００９０】

入出力コントローラ５０－１が第Ｉライトメッセージを受けたとき、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行する。

【００９１】

メモリコントローラ４０－ｉは、第Ｉデータに対してフリー状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉ強書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の強・弱書き込み要求メッセージ両方を受け付けることができないライトロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、入出力コントローラ５０－１に宛てて第Ｉ書き込み許可メッセージをネットワーク２に出力する。

メモリコントローラ４０－ｉは、第Ｉデータに対してフリー状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉ弱書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、前記フリー状態情報に代えて、プロセッサや入出力デバイスからの読み出し要求や、前記第Ｉデータに対する他の弱書き込み要求メッセージを受け付けることができない割り込み可ライトロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、入出力コントローラ５０－１に宛てて第Ｉ書き込み許可メッセージをネットワーク２に出力する。

メモリコントローラ４０－ｉは、第Ｉデータに対してライトロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉ強書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、入出力コントローラ５０－１に宛てて第Ｉ不許可メッセージをネットワーク２に出力する。

メモリコントローラ４０－ｉは、第Ｉデータに対して割り込み可ライトロック状態情報あるいはライトロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉ弱書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、入出力コントローラ５０－１に宛てて第Ｉ不許可メッセージをネットワーク２に出力する。

【００９２】

入出力コントローラ５０－１は、この第Ｉ不許可メッセージを受け取ると、第１から第（Ｉ－１）書き込み許可メッセージまでの（Ｉ－１）個の書き込み許可メッセージを既に

受け取っているか否かを検査し、受け取り済みであれば第Ｉ強書き込み要求メッセージを、受け取り済みでなければ第Ｉ弱書き込み要求メッセージを発行する。

メモリコントローラ４０－ｉは、第Ｉデータに対して割り込み可ライトロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉ強書き込み要求メッセージを前記入出力コントローラ５０－１から受け取った場合、前記割り込み可ライトロック状態情報に代えて、リクエストロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、前記ディレクトリに格納された入出力コントローラを特定する情報（ここでは入出力コントローラ５０－２とする）が指す入出力コントローラ５０－２に宛てて第Ｉ再試行要求メッセージをネットワークに２出力する。

第Ｉ再試行要求メッセージを受け取った入出力コントローラ５０－２は、以下に示す再試行処理を行う。

【００９３】

まず、メモリコントローラ４０－ｉに宛てて第Ｉ開放メッセージをネットワーク２に出力する。

次に、入出力コントローラ５０－２が受け取ったライトメッセージの中で、前記第Ｉデータとアドレスを同じくするライトメッセージで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ更新メッセージの発行を停止する。そして、前記第Ｉ開放メッセージの発行後に、メモリコントローラ４０－ｉに宛てて書き込み要求メッセージを発行する。これで再試行処理は終わる。

メモリコントローラ４０－ｉは、前記第Ｉ開放メッセージを受け取ると、リクエストロック状態情報に代えてライトロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、入出力コントローラ５０－１に宛てて第Ｉ書き込み許可メッセージを前記ネットワークに出力する。

また、メモリコントローラ４０－ｉは、第Ｉデータに対してリクエストロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉデータに対する更新メッセージを前記入出力コントローラ５０－２から受け取った場合、主記憶部３０－２のデータの値を更新メッセージで指定される値に更新する。

【００９４】

また、第３実施例の変形として、次のような変形も可能である。

第Ｉ再試行要求メッセージを受け取った入出力コントローラ５０－２が行う再試行処理で、まず、入出力コントローラ５０－２が受け取ったライトメッセージの中で、前記第Ｉデータとアドレスを同じくするライトメッセージで書き込み許可メッセージを受け取り済みのものに関して、まだ更新メッセージを発行していなければ第Ｉ開放メッセージを発行し、その後にメモリコントローラ４０－ｉに宛てて書き込み要求メッセージを発行し処理を終える。また、まだ更新メッセージを発行していればなにもせず処理を終える。

メモリコントローラ４０－ｉは、前記第Ｉ開放メッセージを受け取ると、リクエストロック状態情報に代えてライトロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、入出力コントローラ５０－１に宛てて第Ｉ書き込み許可メッセージを前記ネットワークに出力する。

また、メモリコントローラ４０－ｉは、第Ｉデータに対してリクエストロック状態情報が前記ディレクトリ２０－ｉに格納されているときに、第Ｉデータに対する更新メッセージを前記入出力コントローラ５０－２から受け取った場合、主記憶部３０－２のデータの値を更新メッセージで指定される値に更新する。そして、リクエストロック状態情報に代えてライトロック状態情報を前記ディレクトリ２０－ｉに格納する。そして、入出力コントローラ５０－１に宛てて第Ｉ書き込み許可メッセージを前記ネットワークに出力する。

【００９５】

以上のように構成、動作することで、実施例２と同じようにデッドロックの危険性を回避することができる。

【図面の簡単な説明】

【００９６】

【図１】従来のマルチプロセッサシステムの構成を示す図である。
【図２】従来技術１の動作を示すタイミングチャート図である。
【図３】従来技術２の動作を示すタイミングチャート図である。
【図４】本発明のマルチプロセッサシステムの構成を示す図である。
【図５】本発明のマルチプロセッサシステムにおける入出力コントローラの構成を示す図である。
【図６】本発明のマルチプロセッサシステムの動作として、ライトメッセージが連続して発行された場合の動作を示すタイミングチャート図である。
【図７Ａ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｂ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｃ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｄ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｅ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｆ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｇ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｈ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図７Ｉ】本発明のマルチプロセッサシステムにおける入出力コントローラの動作例を説明するための図である。
【図８】本発明のマルチプロセッサシステムの他の構成を示す図である。
【図９】本発明の実施例１に係るマルチプロセッサシステムの動作として、ディレクトリの状態がＵｎｃａｃｈｅｄであった場合の動作を示すタイミングチャート図である。
【図１０】本発明の実施例１に係るマルチプロセッサシステムの動作として、ディレクトリの状態がＣｌｅａｎであった場合の動作を示すタイミングチャート図である。
【図１１】本発明の実施例１に係るマルチプロセッサシステムの動作として、ディレクトリの状態がＤｉｒｔｙであった場合の動作を示すタイミングチャート図である。
【図１２】本発明の実施例１に係るマルチプロセッサシステムの動作として、ディレクトリの状態がリクエストロック状態あるいはライトロック状態であった場合の動作を示すタイミングチャート図である。
【図１３】本発明の実施例２に係るマルチプロセッサシステムにおける入出力コントローラの構成を示す図である。
【図１４】本発明のマルチプロセッサシステムにおける入出力コントローラが、第Ｉ書き込み許可メッセージを受けた場合の動作を示すフローチャート図である。

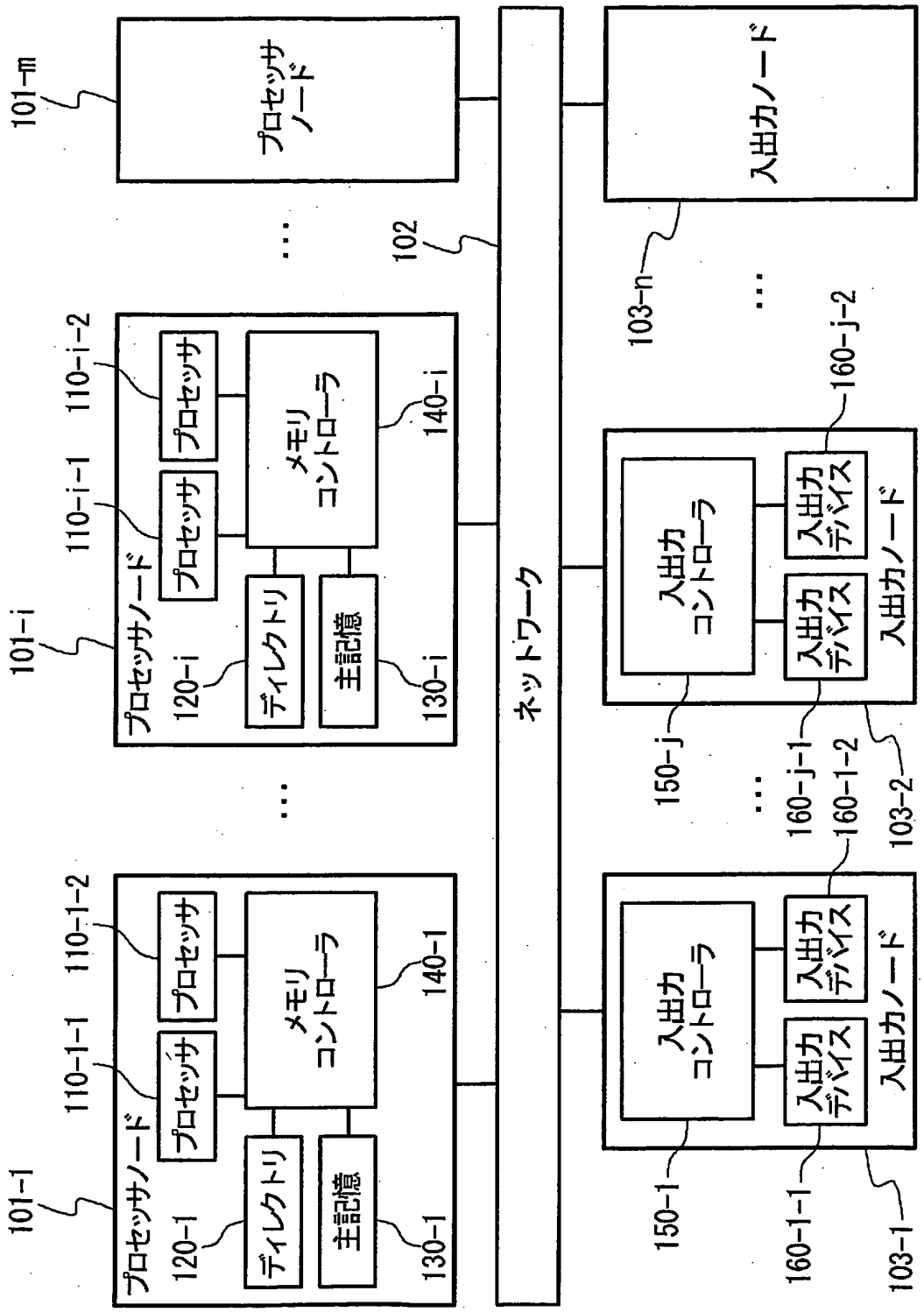
【符号の説明】

【００９７】

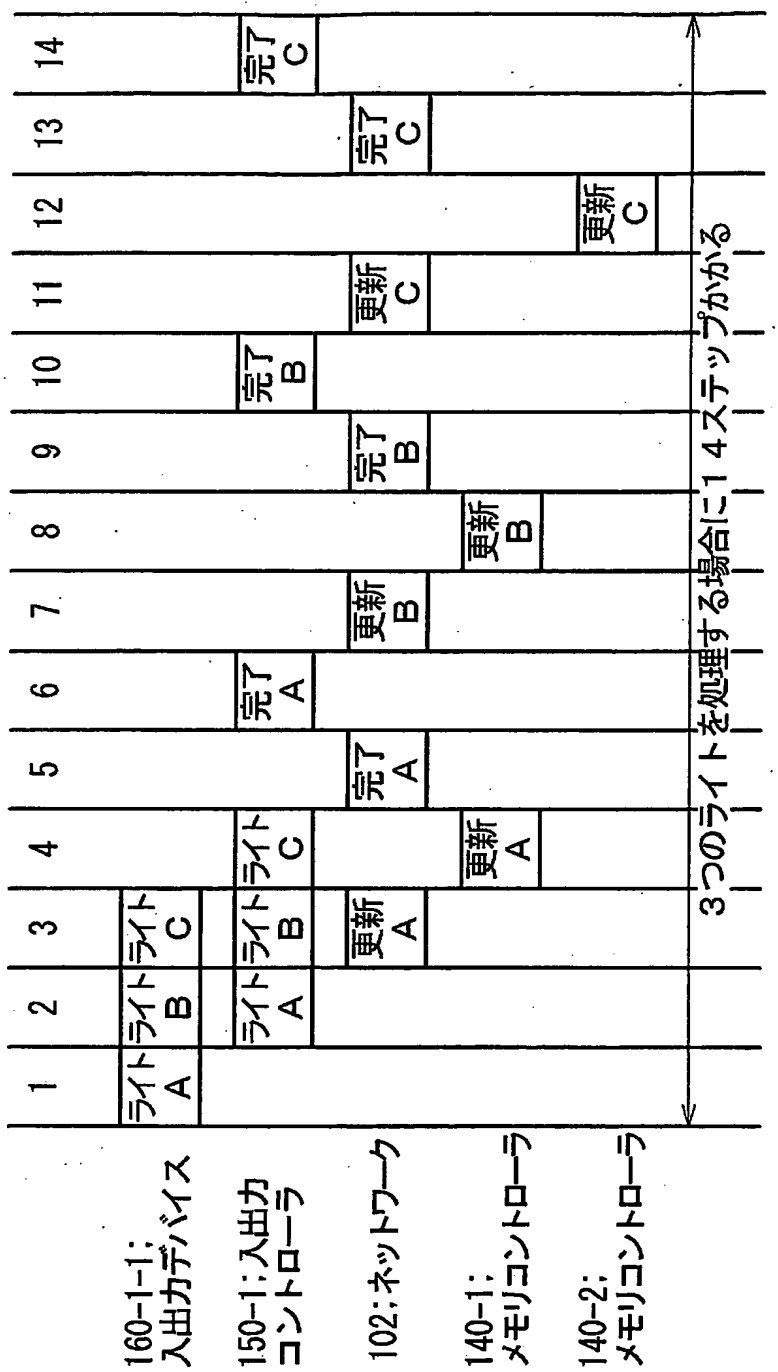
１－１～１－ｍ　　プロセッサノード
２　　ネットワーク
３－１～３－ｎ　　入出力ノード
１０－ｉ－１、１０－ｉ－２（ｉ＝１、２、…、ｍ）　　プロセッサ
２０－ｉ（ｉ＝１、２、…、ｍ）　　ディレクトリ
３０－ｉ（ｉ＝１、２、…、ｍ）　　主記憶
４０－ｉ（ｉ＝１、２、…、ｍ）　　メモリコントローラ

５０－ｊ（ｊ＝１、２、…、ｎ）　入出力コントローラ
５１－ｊ（ｊ＝１、２、…、ｎ）　入出力セレクタ
５２－ｊ（ｊ＝１、２、…、ｎ）　入出力コントローラ
６０－ｊ－１、６０－ｊ－２（ｊ＝１、２、…、ｎ）　入出力デバイス
７１－ｊ（ｊ＝１、２、…、ｎ）　セレクタ
７２－ｊ（ｊ＝１、２、…、ｎ）　メッセージ格納キュー
７３－ｊ（ｊ＝１、２、…、ｎ）　ライトポインタ
７４－ｊ（ｊ＝１、２、…、ｎ）　リードポインタＡ
７５－ｊ（ｊ＝１、２、…、ｎ）　リードポインタＢ
７６－ｊ（ｊ＝１、２、…、ｎ）　許可フラグ
７８－ｊ（ｊ＝１、２、…、ｎ）　開放処理ポインタ
７９－ｊ（ｊ＝１、２、…、ｎ）　開放処理フラグ
１０１－１～１０１－ｍ　プロセッサノード
１０２　ネットワーク
１０３－１～１０３－ｎ　入出力ノード
１１０－ｉ－１、１１０－ｉ－２（ｉ＝１、２、…、ｍ）　プロセッサ
１２０－ｉ（ｉ＝１、２、…、ｍ）　ディレクトリ
１３０－ｉ（ｉ＝１、２、…、ｍ）　主記憶
１４０－ｉ（ｉ＝１、２、…、ｍ）メモリコントローラ
１５０－ｊ（ｊ＝１、２、…、ｎ）　入出力コントローラ
１６０－ｊ－１、６０－ｊ－２（ｊ＝１、２、…、ｎ）　入出力デバイス

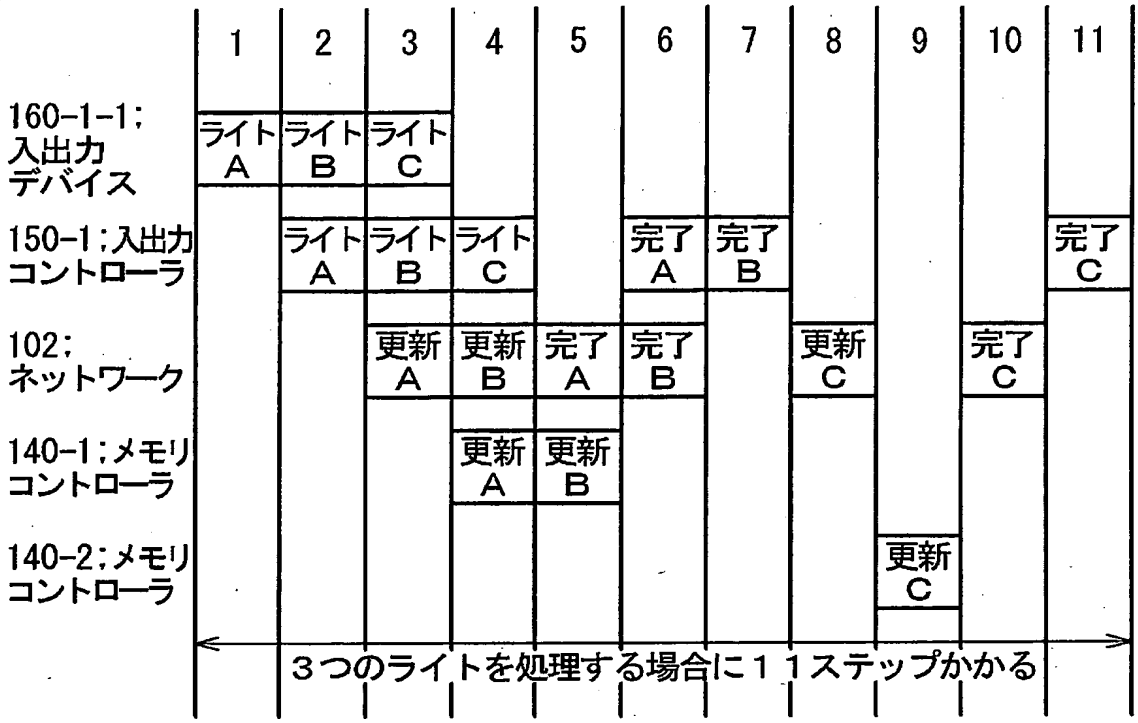
【書類名】図面
【図１】


101-m
プロセッサノード
101-i
プロセッサノード
110-i-2
110-i-1
プロセッサ
プロセッサ
メモリコントローラ
140-i
120-i
ディレクトリ
主記憶
130-i
101-1
プロセッサノード
110-1-2
110-1-1
プロセッサ
プロセッサ
メモリコントローラ
140-1
120-1
ディレクトリ
主記憶
130-1
102
ネットワーク
103-n
入出力ノード
160-j-2
160-j-1
入出力デバイス
入出力デバイス
入出力コントローラ
入出力ノード
150-j
103-2
160-1-2
160-1-1
入出力デバイス
入出力デバイス
入出力コントローラ
入出力ノード
150-1
103-1

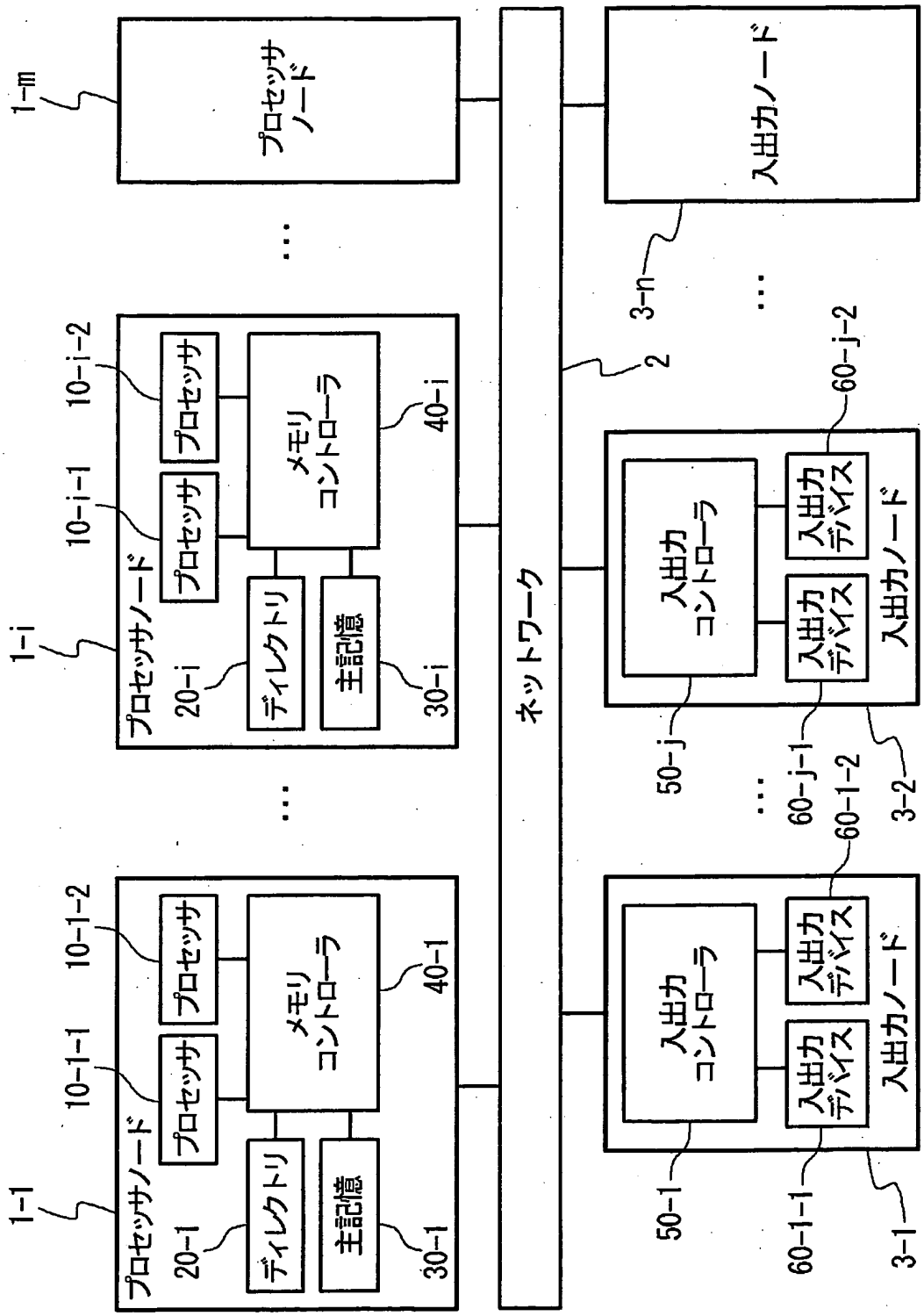
【図２】


1
2
3
4
5
6
7
8
9
10
11
12
13
14
ライトA
ライトB
ライトC
ライトA
ライトB
ライトC
完了A
更新A
完了B
更新B
完了A
更新B
完了B
更新A
完了C
更新C
完了C
更新C
3つのライトを処理する場合に14ステップかかる
160-1-1;
入出力デバイス
150-1;入出力
コントローラ
102;ネットワーク
140-1;
メモリコントローラ
140-2;
メモリコントローラ

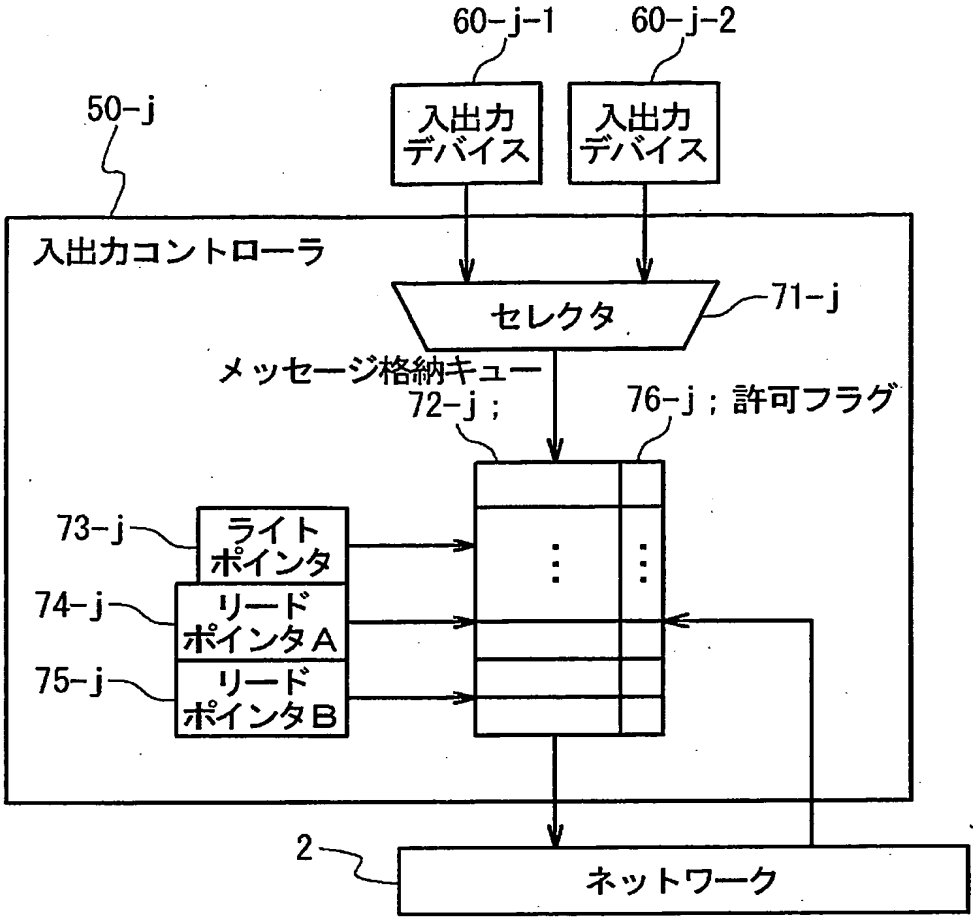
【図３】


1
2
3
4
5
6
7
8
9
10
11
160-1-1;
入出力
デバイス
ライト
A
ライト
B
ライト
C
150-1;入出力
コントローラ
ライト
A
ライト
B
ライト
C
完了
A
完了
B
完了
C
102;
ネットワーク
更新
A
更新
B
完了
A
完了
B
更新
C
完了
C
140-1;メモリ
コントローラ
更新
A
更新
B
140-2;メモリ
コントローラ
更新
C
３つのライトを処理する場合に１１ステップかかる

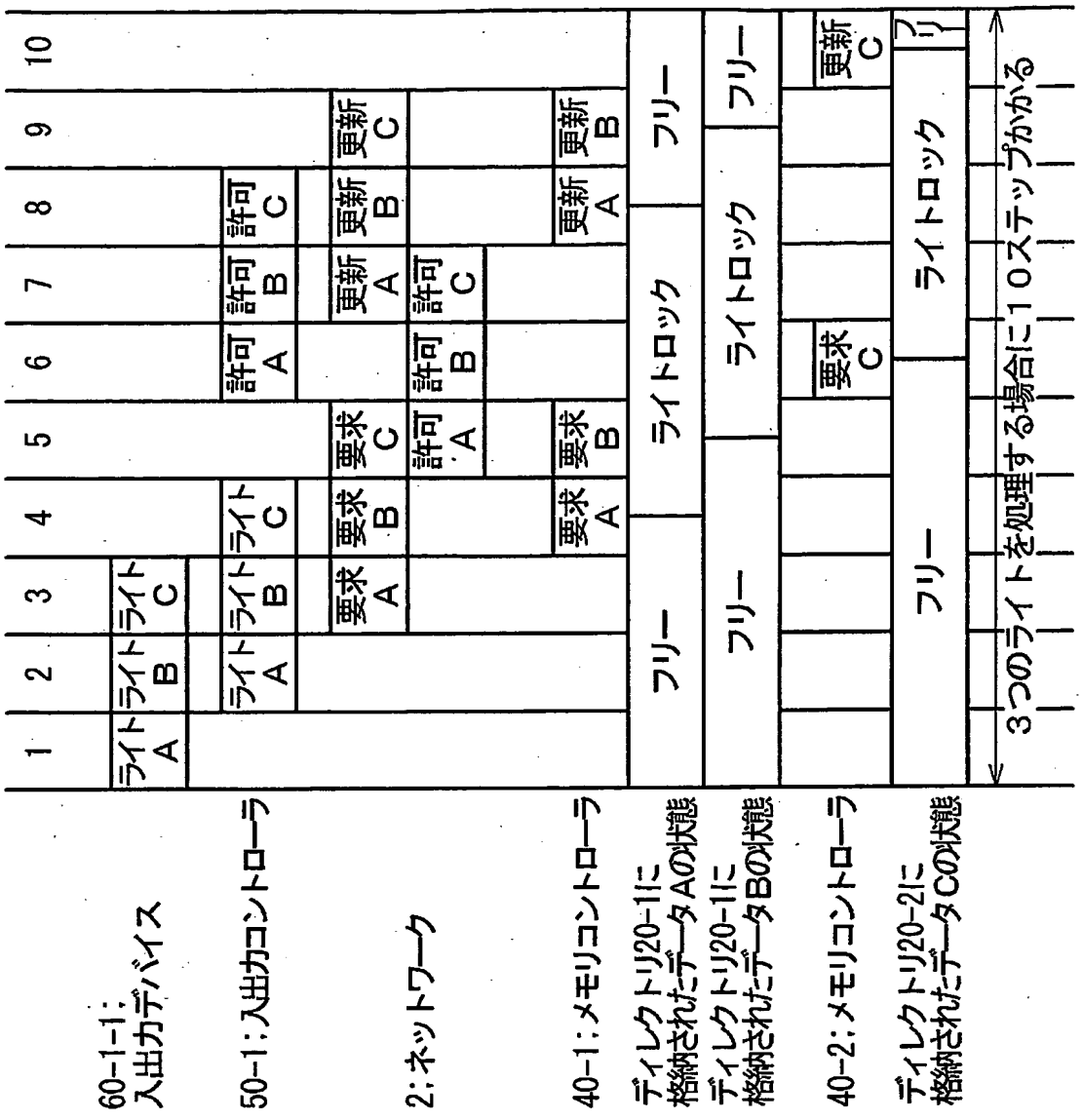
【図４】


プロセッサノード
入出力ノード
ネットワーク
メモリコントローラ
ディレクトリ
主記憶
プロセッサ
入出力コントローラ
入出力デバイス
1-m
1-i
1-1
2
3-n
3-2
3-1
10-i-2
10-i-1
10-1-2
10-1-1
20-i
20-1
30-i
30-1
40-i
40-1
50-j
50-1
60-j-2
60-j-1
60-1-2
60-1-1

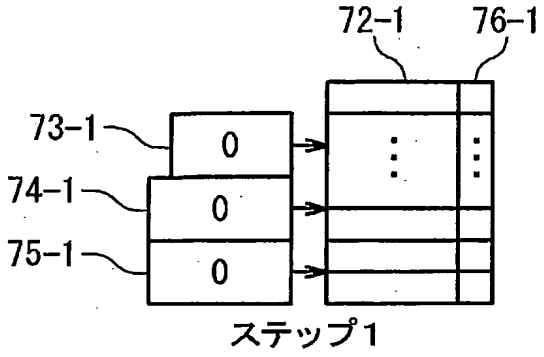
【図５】


60-j-1
60-j-2
50-j
入出力デバイス
入出力デバイス
入出力コントローラ
セレクタ
71-j
メッセージ格納キュー
72-j；
76-j；許可フラグ
73-j
ライトポインタ
74-j
リードポインタA
75-j
リードポインタB
2
ネットワーク

【図６】

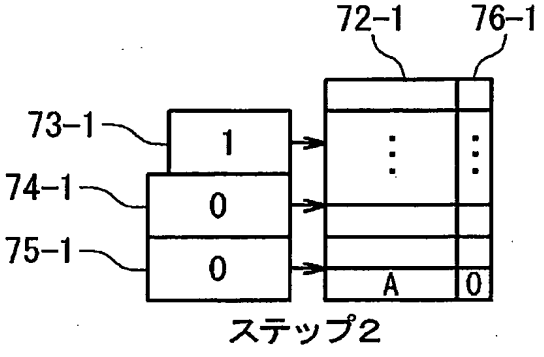

1
2
3
4
5
6
7
8
9
10
ライトA
ライトB
ライトC
ライトA
ライトB
ライトC
要求A
要求B
要求C
許可A
許可B
許可C
更新A
更新B
更新C
要求A
要求B
要求C
許可A
許可B
許可C
更新A
更新B
更新C
フリー
ライトロック
フリー
フリー
ライトロック
フリー
要求C
更新C
フリー
ライトロック
フリー
3つのライトを処理する場合に10ステップかかる
60-1-1：
入出力デバイス
50-1：入出力コントローラ
2：ネットワーク
40-1：メモリコントローラ
ディレクトリ20-1に
格納されたデータAの状態
ディレクトリ20-1に
格納されたデータBの状態
40-2：メモリコントローラ
ディレクトリ20-2に
格納されたデータCの状態

【図７Ａ】



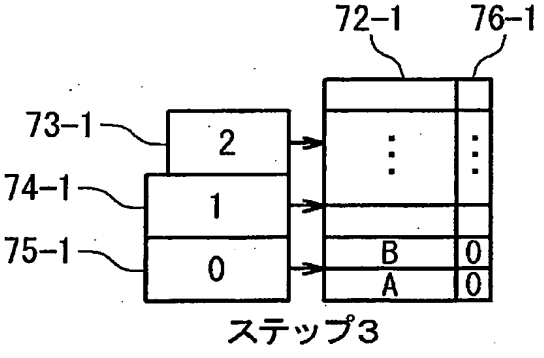
ステップ１

【図７Ｂ】



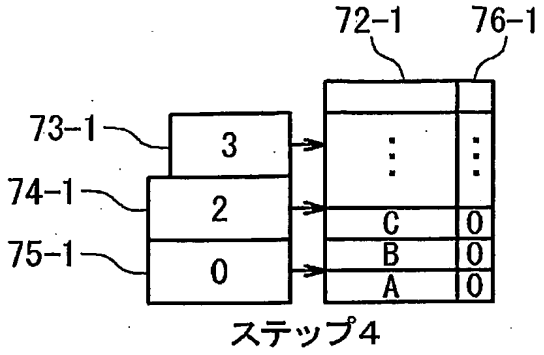
ステップ２

【図７Ｃ】



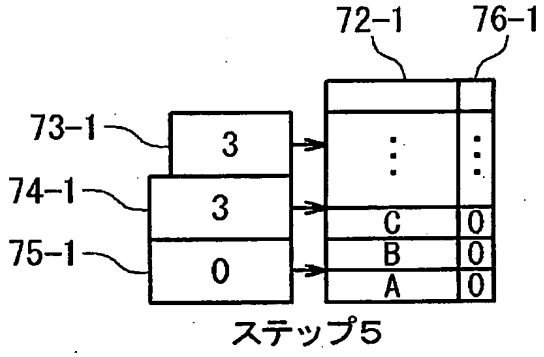
ステップ３

【図７Ｄ】



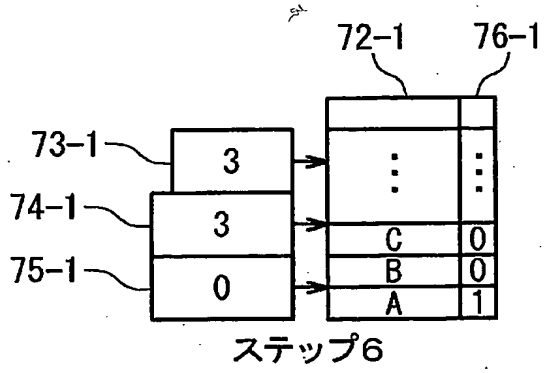
ステップ４

【図７Ｅ】



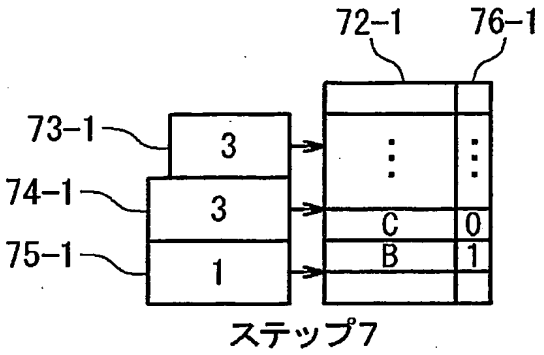
ステップ５

【図７Ｆ】

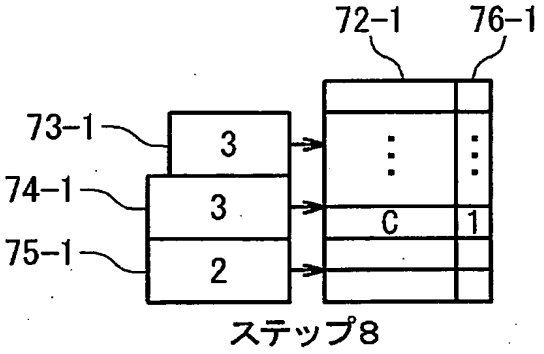


ステップ６

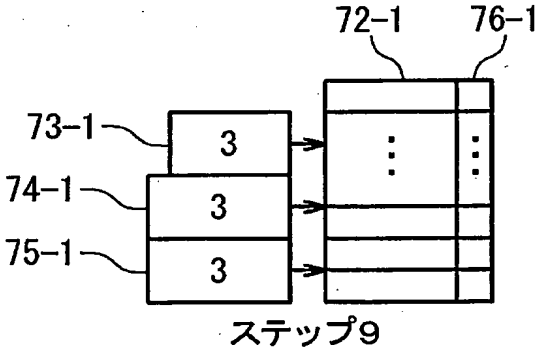
【図７Ｇ】


72-1
76-1
73-1
3
74-1
3
75-1
1
C
0
B
1
ステップ7

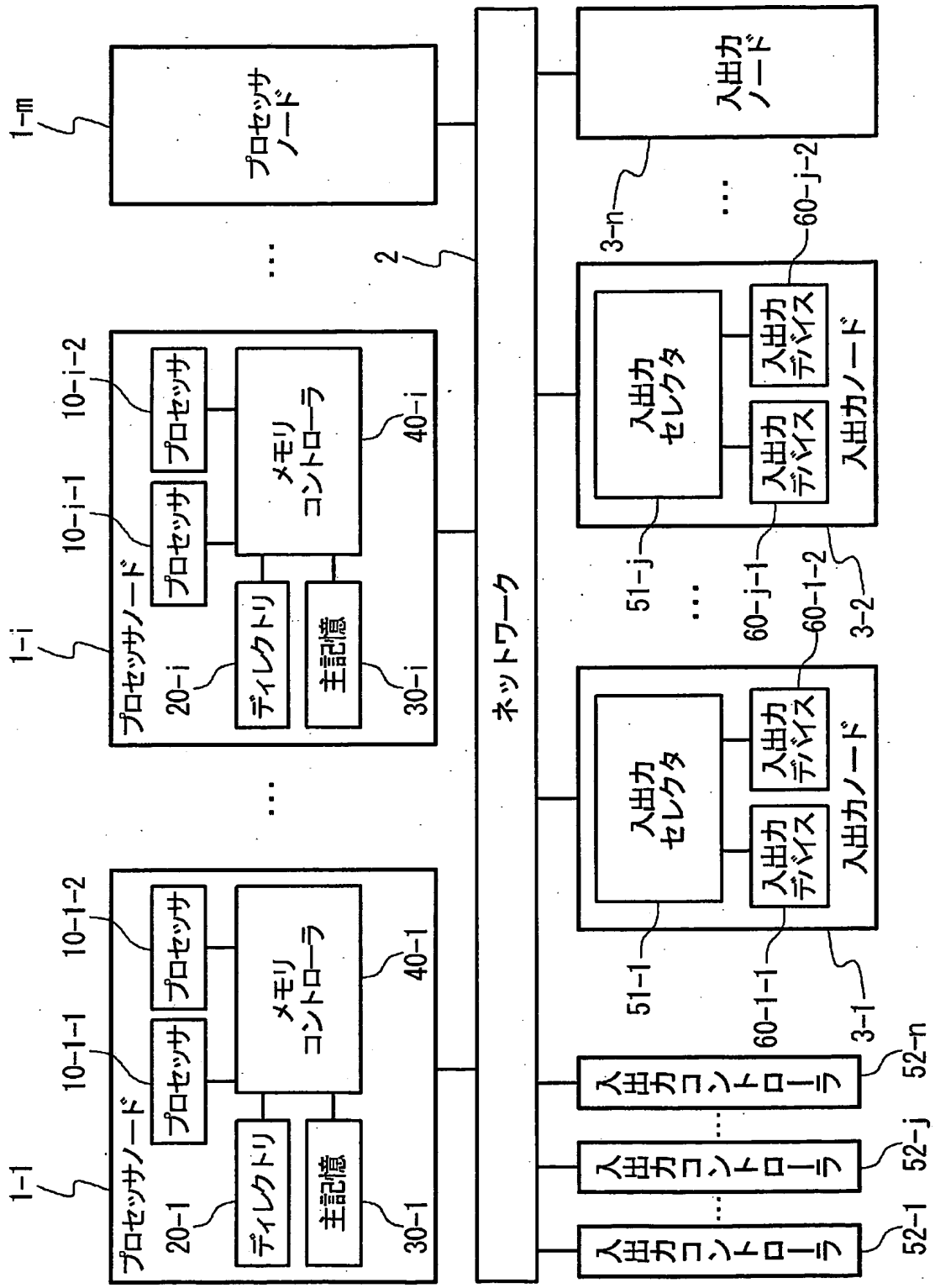
【図７Ｈ】


72-1
76-1
73-1
3
74-1
3
75-1
2
C
1
ステップ8

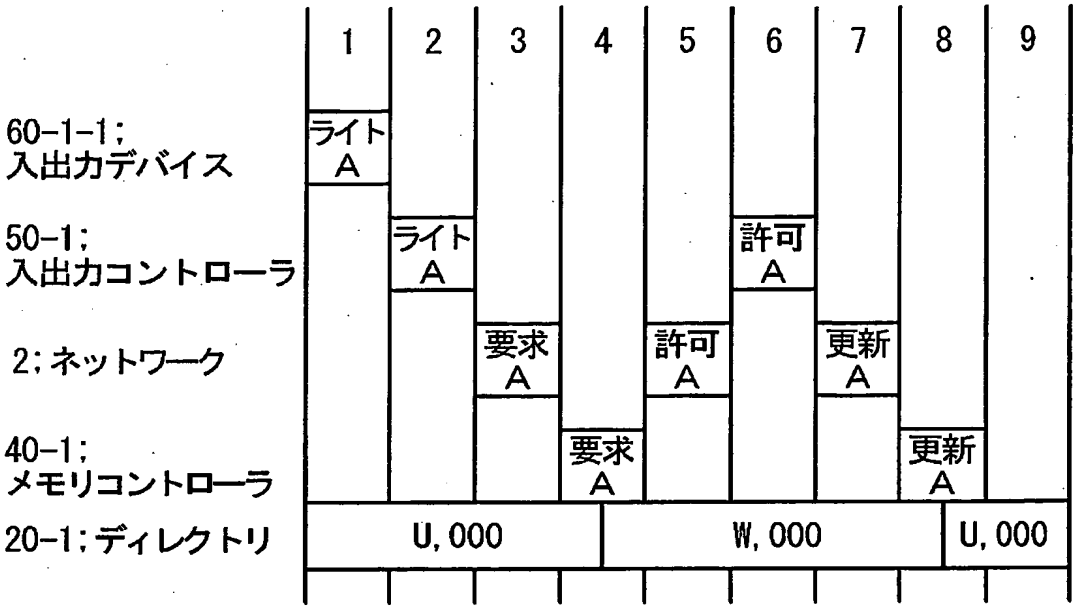
【図７Ｉ】


72-1
76-1
73-1
3
74-1
3
75-1
3
ステップ9

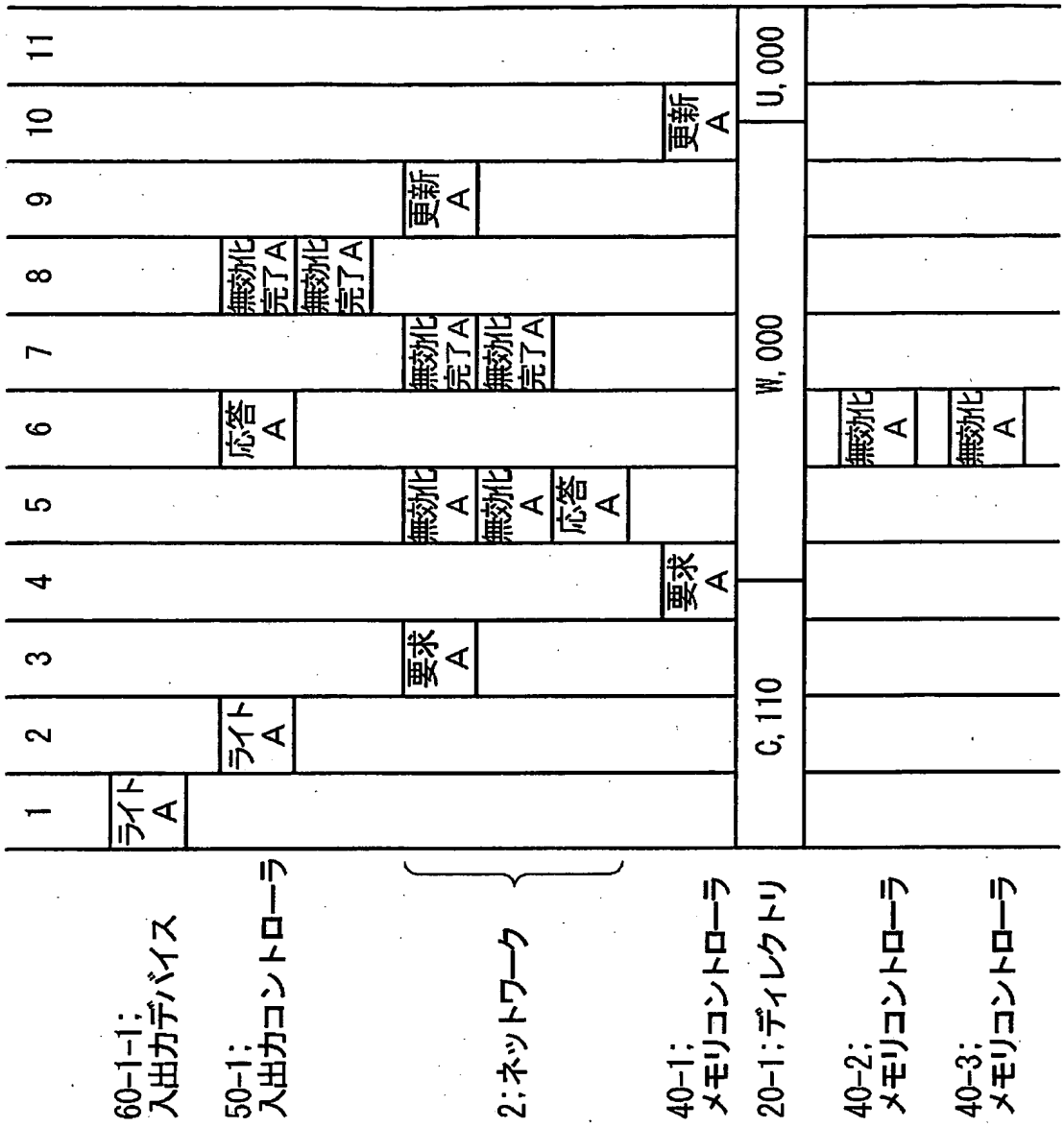
【図８】


1-m
プロセッサノード
1-i
プロセッサノード
10-i-2
プロセッサ
10-i-1
プロセッサ
メモリコントローラ
40-i
20-i
ディレクトリ
主記憶
30-i
1-1
プロセッサノード
10-1-2
プロセッサ
10-1-1
プロセッサ
メモリコントローラ
40-1
20-1
ディレクトリ
主記憶
30-1
2
ネットワーク
3-n
入出力ノード
60-j-2
入出力デバイス
60-j-1
入出力デバイス
入出力セレクタ
51-j
入出力ノード
3-2
60-1-2
入出力デバイス
60-1-1
入出力デバイス
入出力セレクタ
51-1
入出力ノード
3-1
入出力コントローラ
52-n
入出力コントローラ
52-j
入出力コントローラ
52-1

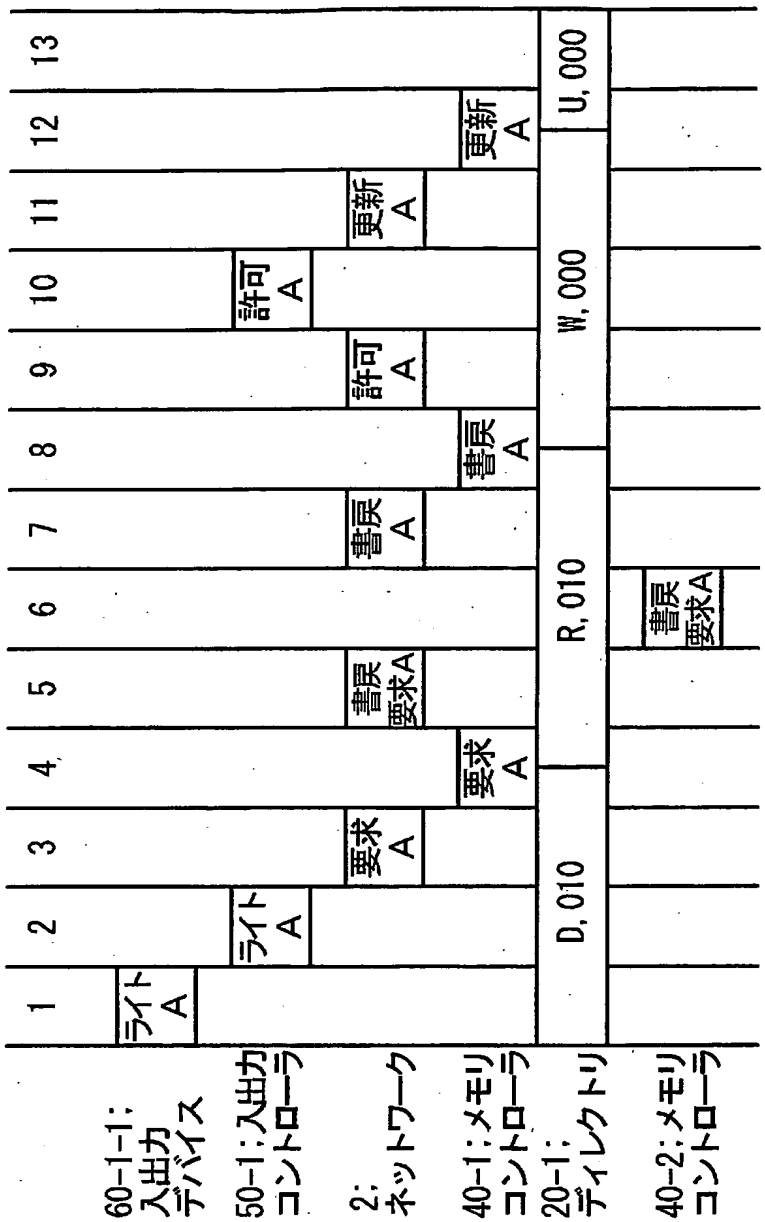
【図９】


1
2
3
4
5
6
7
8
9
60-1-1;
入出力デバイス
50-1;
入出力コントローラ
2;ネットワーク
40-1;
メモリコントローラ
20-1;ディレクトリ
ライト
A
ライト
A
要求
A
要求
A
許可
A
許可
A
更新
A
更新
A
U,000
W,000
U,000

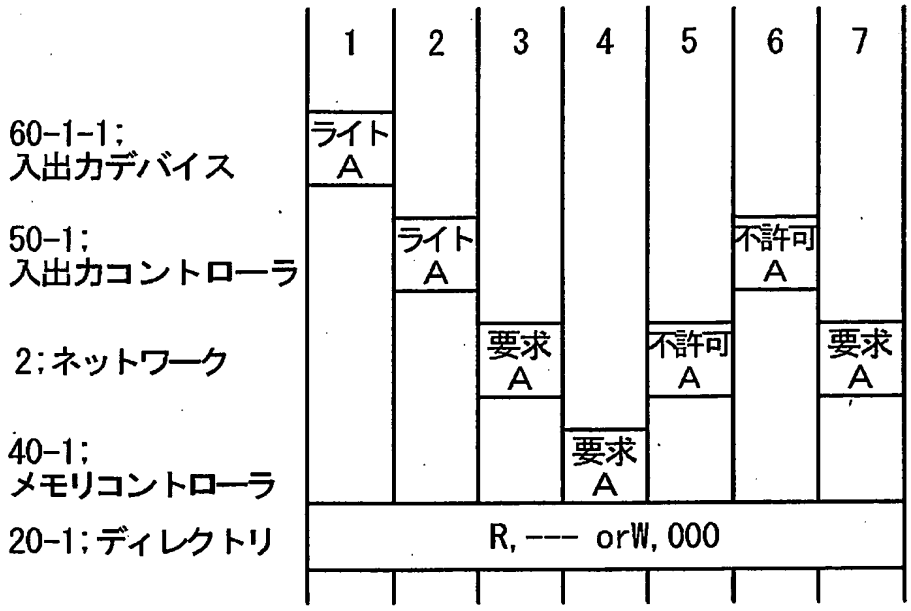
【図１０】


1
2
3
4
5
6
7
8
9
10
11
ライトA
ライトA
要求A
要求A
無効化A
無効化A
応答A
応答A
無効化完了A
無効化完了A
無効化完了A
無効化完了A
更新A
更新A
C, 110
W, 000
U, 000
無効化A
無効化A
60-1-1;
入出力デバイス
50-1;
入出力コントローラ
2;ネットワーク
40-1;
メモリコントローラ
20-1;ディレクトリ
40-2;
メモリコントローラ
40-3;
メモリコントローラ

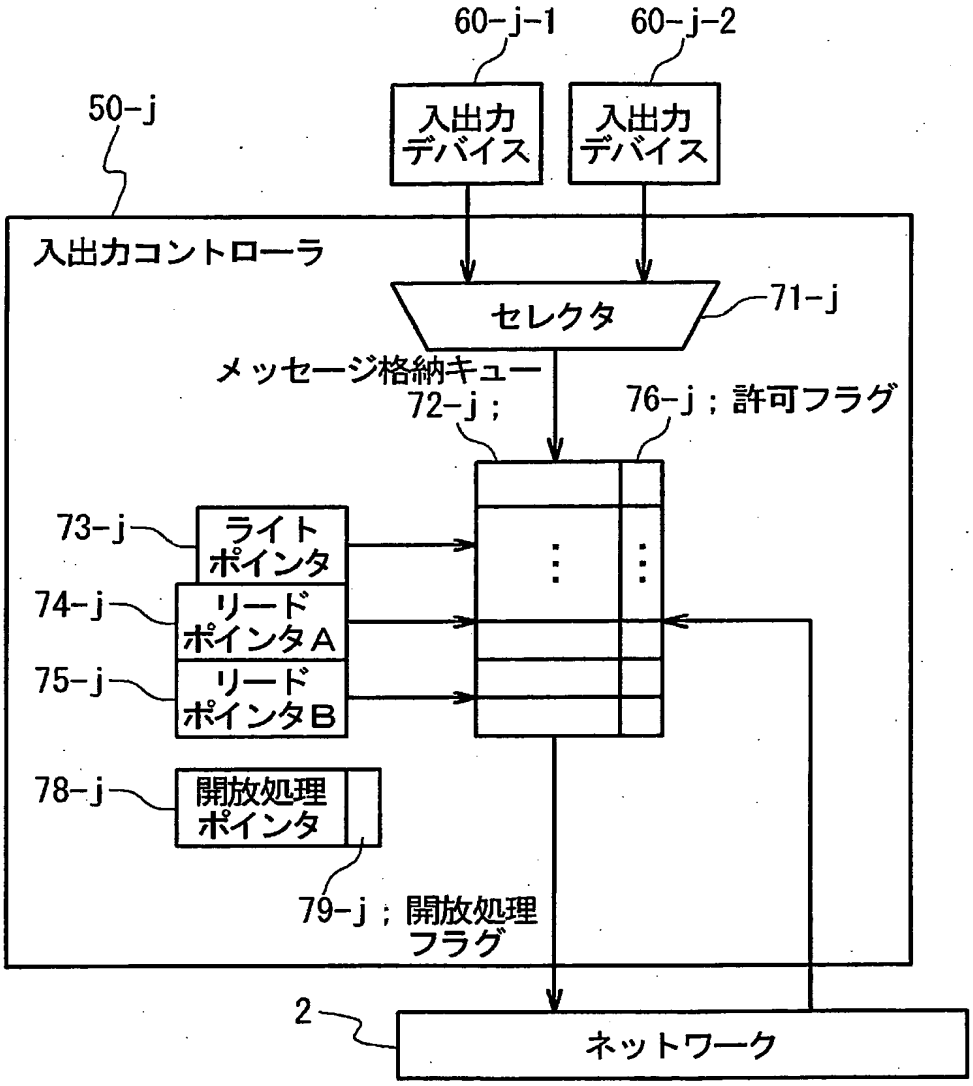
【図１１】


1
2
3
4
5
6
7
8
9
10
11
12
13
ライトA
ライトA
要求A
要求A
書戻要求A
書戻A
書戻A
許可A
許可A
更新A
更新A
D, 010
R, 010
W, 000
U, 000
書戻要求A
60-1-1：入出力デバイス
50-1：入出力コントローラ
2：ネットワーク
40-1：メモリコントローラ
20-1：ディレクトリ
40-2：メモリコントローラ

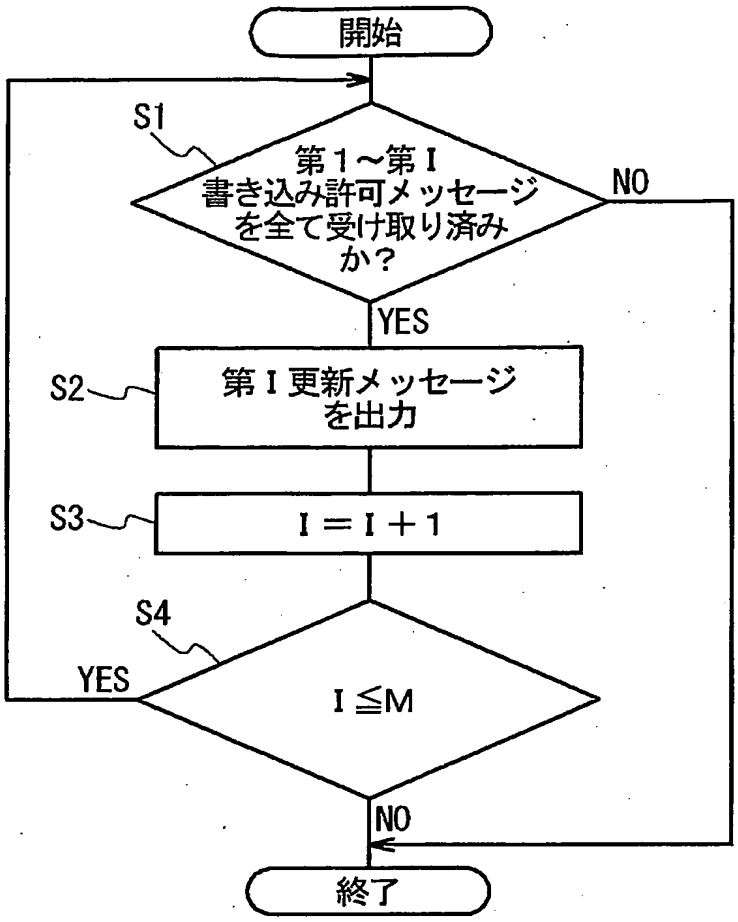
【図１２】


1
2
3
4
5
6
7
60-1-1;
入出力デバイス
50-1;
入出力コントローラ
2;ネットワーク
40-1;
メモリコントローラ
20-1;ディレクトリ
ライト
A
ライト
A
要求
A
要求
A
不許可
A
不許可
A
要求
A
R, --- orW, 000

【図１３】


60-j-1
60-j-2
50-j
入出力デバイス
入出力デバイス
入出力コントローラ
セレクタ
71-j
メッセージ格納キュー
72-j；
76-j；許可フラグ
73-j
ライトポインタ
74-j
リードポインタA
75-j
リードポインタB
78-j
開放処理ポインタ
79-j；開放処理フラグ
2
ネットワーク

【図１４】


開始
S1
第１～第Ｉ
書き込み許可メッセージ
を全て受け取り済み
か？
NO
YES
S2
第Ｉ更新メッセージ
を出力
S3
Ｉ＝Ｉ＋１
S4
YES
Ｉ≦M
NO
終了

【書類名】要約書

【要約】

【課題】入出力コントローラが、異なるプロセッサノードを宛先とする入出力デバイスからの複数のライトメッセージを連続して処理することができるマルチプロセッサシステムを提供すること。

【解決手段】入出力コントローラがライトメッセージを受けたときに、該当するデータをメモリに保持するホームプロセッサノードに書き込み要求メッセージを発行する。書き込み要求メッセージを受け取ったプロセッサノードのメモリコントローラは、ディレクトリに格納された該当するデータの状態に基づいて一貫性処理を行い、該書き込み要求メッセージを発行した入出力コントローラに書込みの許可を示すメッセージが届くように制御する。書き込みの許可を表すメッセージを受け取った入出力ノードの入出力コントローラは、データの書込みを行うライトメッセージとして更新メッセージをホームのプロセッサノードに発行する。更新メッセージを受け取ったプロセッサノードのメモリコントローラは、主記憶部のデータを更新する。上記処理で、入出力コントローラは、入出力デバイスから複数のライトメッセージを受け取ったときに、先行するライトメッセージの進捗にかかわりなく書き込み要求メッセージを発行し、先行するライトのライトメッセージ発行が行われた後にライトメッセージを発行する。

【選択図】図４

特願２００４－１９７２９６

出　願　人　履　歴　情　報

識別番号　　[０００００４２３７]

１．変更年月日　　１９９０年　８月２９日
[変更理由]　　新規登録
住　所　　東京都港区芝五丁目７番１号
氏　名　　日本電気株式会社

特願２００４－１９７２９６

出　願　人　履　歴　情　報

識別番号　　[０００１６８２８５]

１．変更年月日　　２００２年　７月３０日
[変更理由]　　名称変更
住所変更
住　所　　山梨県甲府市大津町１０８８－３
氏　名　　エヌイーシーコンピュータテクノ株式会社